

us-09-486-167a-l.rge

Mon Oct 22 14:46:51 2001

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: October 22, 2001, 10:42:40 ; Search time 1293.29 Seconds  
(without alignments)  
9627.800 Million cell updates/sec

Title: US-09-486-167A-1  
Perfect score: 805  
Sequence: 1 gccaggagcgagtggaag.....ttgtgtttgcggaaaaaaa 805

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues 2688314  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_bal.\*  
2: gb\_ba2.\*  
3: gb\_ba3.\*  
4: gb\_in1.\*  
5: gb\_in2.\*  
6: gb\_in3.\*  
7: gb\_om.\*  
8: gb\_ov.\*  
9: gb\_pat1.\*  
10: gb\_pat2.\*  
11: gb\_ph.\*  
12: gb\_pl1.\*  
13: gb\_pl2.\*  
14: gb\_pl3.\*  
15: gb\_pl4.\*  
16: em\_bal.\*  
17: em\_ba2.\*  
18: em\_fun.\*  
19: em\_htgo\_hum.\*  
20: em\_htgo\_inv.\*  
21: em\_htgo\_rod.\*  
22: em\_htg\_hum1.\*  
23: em\_htg\_hum2.\*  
24: em\_htg\_hum3.\*  
25: em\_htg\_hum4.\*  
26: em\_htg\_hum5.\*  
27: em\_htg\_hum6.\*  
28: em\_htg\_hum7.\*  
29: em\_htg\_hum8.\*  
30: em\_htg\_inv1.\*  
31: em\_htg\_inv2.\*  
32: em\_htg\_other.\*  
33: em\_htg\_rod.\*  
34: em\_hum1.\*  
35: em\_hum2.\*  
36: em\_hum3.\*  
37: em\_hum4.\*  
38: em\_hum5.\*  
39: em\_hum6.\*  
40: em\_hum7.\*  
41: em\_in.\*  
42: em\_om.\*  
43: em\_or.\*

44: em\_ov.\*  
45: em\_pat.\*  
46: em\_ph.\*  
47: em\_pl.\*  
48: em\_ro.\*  
49: em\_sts.\*  
50: em\_sv.\*  
51: em\_un.\*  
52: em\_vi.\*  
53: gb\_sts1.\*  
54: gb\_sts2.\*  
55: gb\_sts3.\*  
56: gb\_sy.\*  
57: gb\_un.\*  
58: gb\_vil.\*  
59: gb\_vil2.\*  
60: gb\_htg1.\*  
61: gb\_htg2.\*  
62: gb\_htg3.\*  
63: gb\_htg4.\*  
64: gb\_htg5.\*  
65: gb\_htg6.\*  
66: gb\_htg7.\*  
67: gb\_htg8.\*  
68: gb\_htg9.\*  
69: gb\_htg10.\*  
70: gb\_htg11.\*  
71: gb\_htg12.\*  
72: gb\_htg13.\*  
73: gb\_htg14.\*  
74: gb\_htg15.\*  
75: gb\_htg16.\*  
76: gb\_htg17.\*  
77: gb\_htg18.\*  
78: gb\_htg19.\*  
79: gb\_htg20.\*  
80: gb\_htg21.\*  
81: gb\_htg22.\*  
82: gb\_htg23.\*  
83: gb\_htg24.\*  
84: gb\_htg25.\*  
85: gb\_pr1.\*  
86: gb\_pr2.\*  
87: gb\_pr3.\*  
88: gb\_pr4.\*  
89: gb\_pr5.\*  
90: gb\_pr6.\*  
91: gb\_pr7.\*  
92: gb\_pr8.\*  
93: gb\_pr9.\*  
94: gb\_rol.\*  
95: gb\_rol2.\*  
96: gb\_in4.\*  
97: gb\_pr10.\*  
98: em\_ba3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	805	100.0	805	9 A99097	A99097 Sequence 1
2	805	100.0	805	88 AF110731	AF110731 Homo sapi
3	800.2	99.4	840	89 AF124993	AF124993 Homo sapi
4	788.8	98.0	875	89 AF242525	AF242525 Homo sapi
5	778.4	96.7	788	89 AF231705	AF231705 Homo sapi
6	759.8	94.4	835	89 AF112212	AF112212 Homo sapi
7	690.2	85.7	780	89 AF197952	AF197952 Homo sapi
8	650	80.7	725	88 AF110736	AF110736 Cercopith

9 648.4 80.5 725 88 AF110734 Papio ham  
10 489 60.7 489 93 HSA249483  
11 462 57.4 604 9 A99105  
12 459.8 57.1 798 7 AF110735  
13 456 56.6 601 9 A99104  
14 442.4 55.0 837 94 AF110732  
15 441.6 54.9 881 7 AF305564  
16 432.8 53.8 675 9 A99101  
17 432.8 53.8 675 9 A99101  
18 431.8 53.8 675 94 AF110733  
19 423.6 53.1 780 9 A99099  
20 423.6 52.6 895 94 AF124994  
21 300.8 37.4 469 9 A99103  
22 287.6 35.7 195316 71 AC044918  
23 284.4 35.3 127234 78 AL139155  
24 275.6 34.2 538 97 HS082615  
25 267.6 33.2 569 97 HS082616  
26 225 28.0 49259 78 AC090391  
27 225 28.0 66266 82 AP000573  
28 225 28.0 151747 82 AP001453  
29 225 28.0 168475 63 AC015698  
30 225 28.0 174562 60 AC005848  
31 213 26.5 157974 82 AP001380  
32 197 24.5 2710 9 A99106  
33 197 24.5 169612 70 AC027393  
34 197 24.5 172112 61 AC010997  
35 195.4 24.3 206148 80 AL390034  
36 193 24.0 62868 64 AC016799  
37 171.6 21.3 177416 62 AC011998  
38 159.4 19.8 2056 7 S65367  
39 154.8 19.2 103152 77 AC087899\_3  
40 154.4 19.2 1869 94 AB041573  
41 147 18.3 195316 71 AC044918  
42 145 13.0 2360 94 AF20872952  
43 92.8 11.5 12299 1 AE006000  
44 90.4 11.2 32270 63 AC014470  
45 90.4 11.2 151611 4 AC009462

ALIGNMENTS

RESULT 1  
LOCUS A99097  
DEFINITION Sequence 1 from Patent WO9909054.  
ACCESSION A99097  
VERSION A99097.1  
KEYWORDS GI:6782047  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 805)  
AUTHORS Falmagne, P., Wattiez, R., Bernard, A., Hermans, C. and Knoops, B.  
TITLE Peroxisome-associated polypeptide, nucleotide sequence encoding said polypeptide and their uses in the diagnosis and/or the treatment of lung injuries and diseases, and of oxidative stress-related disorders  
JOURNAL Patent: WO 9909054-A 1 25-FEB-1999;  
UNIV MONS HAINAUT (BE); FALMAGNE PAUL (BE); WATTIEZ RUDDY (BE);  
BERNARD ALFRED (BE); HERMANS CEDRIC (BE); KNOOPS BERNARD (BE);  
UNIV LOUVAIN (BE)  
FEATURES  
source Location/Qualifiers  
1..805  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
193..681  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAB70589.1"  
/db\_xref="GI:6782048"  
/translation="MAPIKVGDAIPAVEFEGPGNKNVLAELFKKGVLFVPGAF"

TPGSKTHLPGEVQEAELKAKGVQVVAELSLVNDFAVTGEWGRAHKAEGKVRLLADPT  
GAFGETDLDDSLVSLFNGNRRLKRFEMVVDGIVKALNVDFDTGLTCSLAPNIIS  
QL" 200 c 276 g 166 t  
BASE COUNT 163 a 200 c 276 g 166 t  
ORIGIN  
Query Match 100.0%; Score 805; DB 9; Length 805;  
Best Local Similarity 100.0%; Pred. No. 1.8e-193;  
Matches 805; Conservative 0; Mismatches 0; Indels 0; Gaps 0:  
QY 1 gccagagagcggagtggaagtgccgtggcggtggtatggactagctggtgctgctgccc 60  
Db 1 GCCAGAGGCGGAGTGGAAAGTGGCGGTGGCGGTATGGAGCTAGCTGCGGTGGCGCC 60  
QY 61 ctgagacgtcagcggtctatactactgctggtggcggtggcggtgctgctgctgctgccc 120  
Db 61 CTGAGACGCTCAGCGGCTATATCTACTGCTGGTGGCGGTGGCGGTGAGTCTGCGCAGCG 120  
QY 121 gcaacaagacggtgctgagtgagagtggtgctggtggcggtggcggtgctgctgctgccc 180  
Db 121 GCACCAAGACGGTGCAGTGAAGGAGAGTGGCGGTCTGCGGGTCCGCGTTCAGCAGA 180  
QY 181 gccgtgcagccatgcccacatcaagtggtggtgagatgcatccacagcagtgagtggttt 240  
Db 181 GCCGCTGCAGCCATGCCCCATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTT 240  
QY 241 gaagggagcggaggaacaaagtggaacctggcagctgttcaaggcgcaagaaggtgtg 300  
Db 241 GAAGGGACCCAGGGAACAAAGTGAACCTGCAGAGCTGTCAAGGGCAAGAGGGGTG 300  
QY 301 ctgtttgagttctggggccttccacctggtatgttccaaagacacacctgccaggttt 360  
Db 301 CTGTTTGAAGTTCCTGGGGCTTCCACCTCGGATGTTTCCAAAGACACACCTGCCAGG 360  
QY 361 gtggagcaggtgagctctgaaagcgaagtggtgagtggtggtggtggtggtggtggtt 420  
Db 361 GTGGAGCAGGTGAGGCTCTGAAGGCCAAGGAGTCCAGGTGGTGGCTGTCTGAGTGT 420  
QY 421 aatgatgcttggctgagtggtggtggtggtggtggtggtggtggtggtggtggtggtt 480  
Db 421 AATGATGCTTGTGTGAGTGGCGAGTGGGGCCGAGCCACACAGCGGGAAGGAGTTCG 480  
QY 481 ctctggtctatccactggggtggtggtggtggtggtggtggtggtggtggtggtggtt 540  
Db 481 CTCTGCTGATCCACTGGGGCTTGGGAAGGAGAGAGACTTATTACTAGATGATTCG 540  
QY 541 ctggttccatcttgggaatgcagctctcaagaggttctccatggtggtggtggtggtt 600  
Db 541 CTGTTGTCATCTTTGGGAATCGAGTCTCAAGAGGTCTCCATGGTGGTACAGGATGC 600  
QY 601 atagtgaagcgcctgaatgtggaaccagatggcagcagcctcactgagcctgagcctg 660  
Db 601 ATAGTGAAGGCCTCAATGTGGAACAGATGGCAGCGCTCACTGACGCTGGCACCC 660  
QY 661 aatatactcactgctgagcctgagcctgagcctgagcctgagcctcactgagcctt 720  
Db 661 AATATCATCTCAGCTCTGAGGCGCTGGGCGCTGAGATTAATCTCCACCCCTCTATCT 720  
QY 721 cacttgccagcctgctgctggtggtggtggtggtggtggtggtggtggtggtt 780  
Db 721 CACTTGCCAGCCTGTGCTGGGGCTTGCATTTGAATGATTTGGCCAGATTTCTCAATA 780  
QY 781 aacacttggttggcggaataaaaaa 805  
Db 781 AACACTTGTGTTGGGGAATAAAAA 805  
RESULT 2  
LOCUS AF110731  
DEFINITION Homo sapiens antioxidant enzyme B166 mRNA, complete cds.  
ACCESSION AF110731

[illegible]



```

QY 554 ttggaatcagactctcaagagttctccatgtgtgtacagagatggcactagtgaaagccc 613
Db 553 TTGGGAATCAGCTCTCAAGAGTCTTCCATGTGTGTACAGGATGCATAGTGAAGGCC 612
QY 614 tgaatggaacacagatggcaagagctcaactgagcctggcaccacataatcatctac 673
Db 613 TGAATGTGGAACAGATGGACAGGCTTCACTTCAAGAGGCTGGCACCCTATATCATCTCAC 672
QY 674 agctctgagcctgggagagattactctccacccctctctctcactctccagcc 733
Db 673 AGCTCTGAGGCTGGGCCAGATTACTTCTCCACCCCTCCCTATCTCACTTCCAGCC 732
QY 734 ctgtgtgggcccctgcaattggaatgtggccagatcttgcataaacaacttgggtt 793
Db 733 CTGTGCTGGGCCCCTGCAATTGGAATGTGGCCAGATTTCTGCAATAAACACTTGTGGTT 792
QY 794 tgcggaaaaaaa 805
Db 793 TGCGGCCAAAAA 804

RESULT 5
AF231705 AF231705 788 bp mRNA PRI 27-JUN-2000
LOCUS Homo sapiens Alu co-repressor 1 (ACR1) mRNA, complete cds.
DEFINITION AF231705
ACCESSION AF231705
VERSION AF231705.1 GI:8745393
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 788)
Kropotov, A., Sedova, V., Ivanov, V., Sazeeva, N., Tomilin, A.,
Krutilina, R., Oei, S.L., Griesenbeck, J., Buchlow, G. and Tomilin, N.
A novel human DNA-binding protein with sequence similarity to a
subfamily of redox proteins which is able to repress
RNA-polymerase-III-driven transcription of the Alu-family
retroposons in vitro
JOURNAL Eur. J. Biochem. 260 (2), 336-346 (1999)
MEDLINE 99195471
PUBMED 10095767
REFERENCE 2 (bases 1 to 788)
AUTHORS Kropotov, A.V. and Tomilin, N.V.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-2000) Chromosome Stability, Institute of Cytology
RAS, Tikhoretskii Av 4, St.Petersburg 194064, Russia
FEATURES
source location/Qualifiers
1..788
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
1..788
/gene="ACR1"
20..664
/gene="ACR1"
/codon_start=1
/product="Alu co-repressor 1"
/protein_id="AAF78899.1"
/db_xref="GI:8745394"
/translaton="MGLAGVCAILRRSAGYILVGGAGGQSAANAARRCSGEWASGVR
SF3RAAAMAPIKVDAIPAVEFEGEPGNKNLAELFKKGLVFGVPGATPGCSK
THLPGFVEQALAKAGVQVAVCLSVNDAFVTGEWRAHKAEGKVRLLADPTGAFGE
TDLLDDSLVSIFGNRLKRFMSVVDGIVKALNVEPDGTGLTCSLAPNIISOL"
BASE COUNT 153 a 201 c 268 g 166 t
ORIGIN
Query Match 96.7%; Score 778.4; DB 89; Length 788;
Best Local Similarity 99.9%; Pred. No. 1e-186;
Matches 779; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 19 agtggccctggggggggtatgggactagctggcgtgtgtcgccctgagacgctcagcggc 78
Db 2 AGTGGCCCTGGGGGGGGTATGGGACTAGCTGGCGTGTGGCCCTGAGACGCTCAGCGGGC 61
QY 79 tatatactcgtcgtggggccggtcagctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 138
Db 62 TATATACTCGTTCGTGGGGCCGGCGGTGAGTCTGGCGGAGCGGCGAGCAAGACGGTGCAGT 121
QY 139 gaagagagatgggctcgtggcggggtccgcaggtttccagcagagccgctgcagccatggcc 198
Db 122 GAAGGAGAGTGGCGCTCTGGCGGGGTCCGCAGTTCACGAGAGCGCTGCAGCCATGGCC 181
QY 199 ccaatacgaagtggtggagatgcataccacagctggaggtgtttgaaagggagggaggaac 258
Db 182 CCAATCAAGGTGGGAGATGCCATCCACAGTGGAGGTGTTTGAAGGGGAGGCCAGGGAAC 241
QY 259 aaggtgaacctggcagagctgttcaagggaagaggtgtgtgtgtgtgtgtgtgtgtgtgt 318
Db 242 AAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGAGGTGTGCTGTTTGGAGTTCCTGGG 301
QY 319 gcttcacccctggatgtttccaagacacacacctggccaggggtttgtgtgtgtgtgtgtgt 378
Db 302 GCCTTACCCCTGGATGTTCCAAGACACACCTGGCAGGGTGTGTGAGCAGGCTCAGGCT 361
QY 379 ctgaagggccaaaggtccaggtggtgacctgtctgagtgtaataatgagcctttgtgact 438
Db 362 CTGAAGGCCAAGGAGTCCAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
QY 439 ggcagatggggccgagccacacacacacacacacacacacacacacacacacacacacac 498
Db 422 GCGAGTGGGGCCGAGCCACCAAGGCGAAGGTTCGGCTCCTGGCTGATGCCACT 481
QY 499 ggggctttgggaaggagacagacttattactagatgattcgtgtgtgtgtgtgtgtgtgtgt 558
Db 482 GGGGCTTTGGGAAGGAGACACACTTATTACTAGATGATTTCCCTGCTGTGTCTTCTGGG 541
QY 559 atcgaactcacaagagttccatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 618
Db 542 AATCGACGCTCTCAAGAGGTTCCTCATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 601
QY 619 gtggaaccagatggcacagcctcacctcagcctcagcctcagcctcagcctcagcctcagcct 678
Db 602 GTGGAACACAGATGGCACAGGCTCCTCCTGCGCCCTGGCACCCCAATATCATCTCAGAGTC 661
QY 679 tgaggccctgggcccagattacttctccaccctccctctcactcactcactcactcactcact 738
Db 662 TGAGGCCCTGGCGACAGATTACTTCTCCACCCTCCCTATCTCAGCTGCGCCAGCCCTGTG 721
QY 739 ctggggccctgcaattggaatgttgccagattctgcaataaacaactgtgtgtgtgtgtgt 798
Db 722 CTGGGCCCTGCAATTGGAATGTGGCCAGATTTCTGCAATAAACACTTGTGTGTGTGTGTGT 781

RESULT 6
AF112212 AF112212 835 bp mRNA PRI 13-DEC-1999
LOCUS Homo sapiens putative peroxisomal antioxidant enzyme mRNA, complete
DEFINITION cds.
ACCESSION AF112212
VERSION AF112212.1 GI:6563211
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 835)
Ren, S., Shi, J., Huang, C., Jiang, C., Li, Y., Zhou, J., Yu, Y., Xu, S.,
Wang, Y., Fu, G., Chen, Z. and Han, Z.
A novel gene expressed in human adrenal gland
JOURNAL Unpublished
REFERENCE Ren, S., Shi, J., Huang, C., Jiang, C., Li, Y., Zhou, J., Yu, Y., Xu, S.,
AUTHORS Wang, Y., Fu, G., Chen, Z. and Han, Z.

```

Db	659	TCTCACAGCTCTGAGAGCCCTGGGCCAGATTCTCCCTCCACCCCTCCCTATCTCACCTG	718
Qy	727	ccagccctgtgctggggccctgcaattggaatgttgccagattctgcataaacact	786
Db	719	CCGACCCCTGTGCTGGGGCCCTGCAATTGGGAATGTTGCCAGATTCTGCAATAAACACT	778
Qy	787	tgtggttcgcgaaaaaaa	805
Db	779	TGTGTTTGGCGCCAAAAA	797
RESULT	7		
LOCUS	AF197952	780 bp	mRNA PRI 29-FEB-2000
DEFINITION	Homo sapiens thioredoxin peroxidase PMP20 mRNA, complete cds.		
ACCESSION	AF197952		
VERSION	AF197952.1	GI:6166492	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	1 (bases 1 to 780)		
AUTHORS	Zhou, Y., Kok, K.H., Chun, A.C., Wong, C.M., Wu, H.W., Lin, M.C., Fung, P.C., Kung, H. and Jin, D.Y.		
TITLE	Mouse peroxiredoxin V is a thioredoxin peroxidase that inhibits p53-induced apoptosis		
JOURNAL	Biochem. Biophys. Res. Commun. 268 (3), 921-927 (2000)		
MEDLINE	20145535		
REFERENCE	2 (bases 1 to 780)		
AUTHORS	Jin, D.-Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-OCT-1999) Institute of Molecular Biology, The University of Hong Kong, 8 Sassoon Road, Hong Kong		
FEATURES	Location/Qualifiers		
source	1..780		
CDS	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="11"		
	/map="11q13"		
	110..754		
	/note="type V peroxiredoxin"		
	/codon_start=1		
	/product="thioredoxin peroxidase PMP20"		
	/protein_id="AAF04856.1"		
	/db_xref="GI:6166493"		
	/translation="MGLAGVCAILRRSAGYILVGGAGGQAAAAARRCSEGEWASGVVR SFSSRAAAMAPIKVGDAPAVEFEGEPGNKVNLAELFKGKGLVGVPGAFPGGSK THLPGVDEAEALKGVOVACLVSINDAVFTGEMGRATKAEGKVRLLADPTGAFGSK TDLLEDLSLVISFGNRLRFRFSVVDGIVKALNVEPDGTGLTCSLAPNIISQ"		
BASE COUNT	147 a 203 c 278 g 152 t		
ORIGIN			
Query Match	85.7%;	Score 690.2;	DB 89; Length 780;
Best Local Similarity	99.3%;	Pred. No. 2.2e-164;	
Matches	704; Conservative	0; Mismatches	3; Indels
Gaps	1;		
QY	1	gccagaggcgagtggaagtggccgtggggcggggtatggactagctggcggtgctgcgcc	60
Db	74	GCCAGAGGCGGAGTGGAAGTGGCCGTGGGCGGGTATGGGACTAGTGGCGTGGCGCC	133
QY	61	ctgagacgtcagcgggctatactcgtcgttggggccgcggtcagctgcggcagcg	120
Db	134	CTGAGACGCTACGCGGGCTATATACTCGTGGTGGGCGGGTCACTAGTCTGCGGACGC	193
QY	121	gcagcaagacgtgcagtgaaagagagtggtgcctctgctggcggtccgcagtttcagcaga	180
Db	194	GCAGCAAGACGGTGCAGTGAAGGAGAGTGGGCGGTCTGGCGGGGTCCGACGTTTCAGCAGA	253
QY	181	gcgcgtgcagccattgcccacaatcaaggtggagatgccatccagcagtggaagtgttt	240

Db 254 GCCGCTGAGCCATGCGCCCAATCAAGGTGGAGATGCCATCCAGCAGTGGAGGTGTTT 313  
Qy 241 gaagggagcaggaaggaagtgaaactgagcagagctgttcaaggaggaagaggggtgtg 300  
Db 314 GAAGGGAGCAGGAGACAGGTGAACCTGGCAGAGCTGTTCAGGGCAGAGAGGGTGTG 373  
Qy 301 ctgtttggagtctctggggccttcaaccctggatgtttccaaagacacacacccagggttt 360  
Db 374 CTGTTTGGAGTCTCTGGGGCTTCAACCCCTGGATGTTTCCAAGACACACACCTGCCAGGTTT 433  
Qy 361 gtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420  
Db 434 GTGAGCAGGCTGAGGCTCTGAAGCCCAAGGAGTCCAGGTGGTGGCTCTCTGAGTGT 493  
Qy 421 aatgatgccttggactggcagagtggtggcagccacacagcagcagcagcagcagcagcagc 480  
Db 494 AATGATGCTTGTGACTGGGAGTGGGGCCGAGCCACAAAGGGGAGGCAAGGTTCCG 553  
Qy 481 ctctggctgctccactggggccttgggaagagacagacttattactagatgattcg 540  
Db 554 CTCCTGGCTGATCCACTGGGGCTTGGGAAGGAGACAGACTTATTACTAGATGATTCG 613  
Qy 541 ctgtgtccactcttggaaatcaactctcaagaggttccatgtgtacagagatggc 600  
Db 614 CTGTTGTCATCTTTGGGAATCGAGCTCTCAAGAGGTCTCCATGTTGGTACAGGATGC 673  
Qy 601 atagtgaaggccctgaatgtgaaccagatggcagagcctcactcagcctgagcagcc 660  
Db 674 ATAGTGAAGCCCTGAATGTGGAACAGATGGCAGACGGCCTCAGCTGACGCTGGCACCC 733  
Qy 661 aatatcatctcaacgctctgagggccctggcagagattactctccacc 709  
Db 734 AATATCATCTCACAGCTCTGA--CCCTGGGCCAGATTACTTCTCTCCACC 780

RESULT 8  
AF110736 725 bp mRNA PRI 27-SEP-2000  
LOCUS Cercopithecus aethiops peroxiredoxin 5 (PRDX5) mRNA, complete cds.  
DEFINITION  
ACCESSION AF110736  
VERSION AF110736.2 GI:10305337  
KEYWORDS  
SOURCE African green monkey.  
ORGANISM Cercopithecus aethiops  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Cercopithecoidea;  
Cercopithecoinae; Cercopithecus.  
REFERENCE  
AUTHORS Knoop, B., Clippel, A., Bogard, C., Arsalane, K., Wattiez, R.,  
Hermans, C., Duconsellie, E., Falmagne, P. and Bernard, A.  
TITLE Cloning and characterization of AOEB166, a novel mammalian  
antioxidant enzyme of the peroxiredoxin family  
J. Biol. Chem. 274 (43), 30451-30458 (1999)  
JOURNAL  
MEDLINE 99452929  
PUBMED 10521424  
REFERENCE  
2 (bases 1 to 725)  
Knoop, B. and Cherif, H.  
TITLE Cloning and characterization of COS-7 AOEB166/PRDX5  
Unpublished  
3 (bases 1 to 725)  
Knoop, B., Wattiez, R., Falmagne, P., Hermans, C. and Bernard, A.  
TITLE Direct Submission  
Submitted (04-DEC-1998) Department of Biology, Catholic University  
of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium  
4 (bases 1 to 725)  
Knoop, B., Wattiez, R., Falmagne, P., Hermans, C. and Bernard, A.  
TITLE Direct Submission  
Submitted (26-SEP-2000) Department of Biology, Catholic University  
of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium  
JOURNAL  
REMARK Sequence update by submitter  
COMMENT On Sep 26, 2000 this sequence version replaced gi:10129963.  
FEATURES  
Location/Qualifiers  
1..725  
source

/organism="Cercopithecus aethiops"  
/db\_xref="taxon:9534"  
/cell\_line="COS-7"  
1..725  
/gene="PRDX5"  
5..552  
/gene="PRDX5"  
/note="AOEB166; PRDX5; alkyl hydroperoxide reductase;  
thioredoxin peroxidase; mitochondrial, peroxisomal, and  
cytosolic"  
/codon\_start=1  
/product="peroxiredoxin 5"  
/protein\_id="AAG13453.2"  
/db\_xref="GI:10305338"  
/translation="MGLAGVGLRRSAGYILGGAARQSAVAATAARRRSEGGWASGGV  
KSFRAAAMAPIKVGDAPAVEFVEGEPGNVLAELFKGKGLFVGPQAFTPGGS  
KTHLPFGVEQAEALKAQVQLACLISVNDAEVTEWGRHAKGKRLVLLADPTGAFGK  
ETDLLDLSLVSIFGNRLKRFMSVVDGIYKALNPEDPTGLTCLSLAPSIISQL"  
164  
misc\_feature  
/gene="PRDX5"  
/note="probable alternative translation start site"  
BASE COUNT 141 a 189 c 245 g 150 t  
ORIGIN  
Query Match 80.7%; Score 650; DB 88; Length 725;  
Best Local Similarity 95.3%; Pred. No. 3.3e-154;  
Matches 694; Conservative 0; Mismatches 25; Indels 9; Gaps 2;  
Qy 33 gggatggaactagctggcgtgtggccctgagcagcctcagcagcctatactcgtcgg 92  
Db 1 GGGTATGGGACTAGCTGGCGGTGCTCTGAGACGCTCAGCGGGCTATATAC---TCGG 57  
Qy 93 tggggccggcgggtcagctctgcgcgcgcg-----gcagcaagacagctgcagtgaggaga 146  
Db 58 TGGGGCCGCCCTCAGCTCTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGGAGG 117  
Qy 147 gtggcgctgtgcgggggtccgcagtttcagcagagcgcctgcagccatggccccaatcaa 206  
Db 118 GTGGCGGTCTGGCGGGGTCCGAGTTTCAGCAGAGCGCTGCAGCCATGCCCGCATCAA 177  
Qy 207 ggtggagatgccatcccgagtgagtggtttgaaggaggagcaggaagagtgaa 266  
Db 178 GGTGGAGATGCCATCCCTGCAGTGAGGTGTTGAAGGGGAGCCAGGAGCAAGGTGAA 237  
Qy 267 cctggcagagctgttcaagggaaggggtgtgctgtttgttgagttcctggggcctcac 326  
Db 238 CCTGGCAGAGCTGTTCAAGGGCAAGAGGTGTGCTGTTTGGAGTTCCCGGGCCCTTAC 297  
Qy 327 cctgtgatgttcaagacacacacctgcagggtttgtgagcagcgtcagggctcgaaggc 386  
Db 298 GCCTGGATGTTCCAAAGACCCACCATTACAGGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGC 357  
Qy 387 caaggagtcagggtgtgctgtcgtgagtgatgatgatccttctgactggcagtg 446  
Db 358 CAAAGAGATCCAGGTGTGGCCCTGCTGAGTGTTAATGATGCTTGTGACTGGCCAGTG 417  
Qy 447 ggccgagcccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 506  
Db 418 GGCCGAGCCCAAGGCAAGGTTCGGCTCCTGGCTGATGCCACTGGGGCCCTT 477  
Qy 507 tgggaaggagacagacttattactagatgattcgtggtgtccatctttgggaatcagc 566  
Db 478 TGGGAAGGAGACAGACTTATTACTAGATGATTGCTGATGCTTGTGGAAATCGAGC 537  
Qy 567 tctcaagaggttctccatggtgtacagatggcatagtgaaagcctcgaatgtgaaac 626  
Db 538 TCTCAAGAGGTTCTCCATGGTGTACAGATGGCATAGTGAAGCCCTGAAATGTGGAACC 597  
Qy 627 agatggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 686  
Db 598 AGATGACACAGGCTCACCTGTCAGTCTGGCACCAGCATCATCTCACAGCTCTGAGGCC 657



TITLE Direct Submission  
JOURNAL Submitted (13-SEP-1999) Kim I.H., Biochemistry, Paichai University,  
439-6 Doma-2-Dong Seo-GU, Taejon 302-735, REPUBLIC OF KOREA

FEATURES  
source 1..489  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_line="HeLa Cell"  
gene 1..489  
/gene="TPX VI"  
CDS 1..489  
/gene="TPX VI"  
/function="peroxidase"  
/codon\_start=1  
/evidence="experimental"  
/product="human thiol peroxidase homologous protein"  
/protein\_id="CAB62210.1"  
/db\_xref="GI:6523289"  
/translation="MAPIKVGDAIPAVEVEFEGEPGNKNVLAELFKGKGVLFVGPVGF  
TPGCKTHLPVGEAEALKAKGVQVACLVDYAFVTGEMGRAHKAEGKRVLLADPT  
GAFGKETDLLDDSLVSIFGNRRLRKFSMVVDGIVKALNVEPDGTGLTCSLAPNIIS  
OL"

BASE COUNT 107 a 114 c 163 g 105 t  
ORIGIN

Query Match 60.7%; Score 489; DB 93; Length 489;  
Best Local Similarity 100.0%; Pred. No. 2.1e-113;  
Matches 489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 193 atgcccccaatcaagtgaggatgccatccacagcagtgaggtgtttgaagggagcca 252  
|||||  
Db 1 ATGCCCCCAATCAAGTGGGAGATGCCATCCACAGCAGTGAGGTGTTGAAGGGAGCCA 60  
|||||

Qy 253 ggaacaagtgaaactggcagagctgttcaagggaaggaaggtgtgtgtttgagtt 312  
|||||  
Db 61 GGAACAAGGTGAACCTGGCAGAGCTGTTCAGGGGCAAGAAGGTGTGTGTTGGAGTT 120  
|||||

Qy 313 cctggggccttccccctggatgttccaagacacacctgccagggtttgtgagcaggct 372  
|||||  
Db 121 CCTGGGSCCTTACCCTGGATGTTCCAGACACACCTGCCAGGTTTGTGGAGAGGCT 180  
|||||

Qy 373 gaggtctgaagccaagggagtcacaggtgtgtggccctgtctgagtgtaaatgatgccttt 432  
|||||  
Db 181 GAGGCTGTGAAGGCCAAGGAGTCCAGGTGGTGGCCCTGTCTGAGTGTATGATGCTTT 240  
|||||

Qy 433 gtactgacagtggggcccagcccacaaagcgggaaggaaggttcggctcctggctgat 492  
|||||  
Db 241 GTGACTGGGAGTGGGGCCGAGCCCAAGGGCAAGGTTCCGGCTCCTGGCTGAT 300  
|||||

Qy 493 cccactggggcctttgggaaggagacagacttattactagatgattcgtggtgtccatc 552  
|||||  
Db 301 CCCACTGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCGGTGGTGTCCATC 360  
|||||

Qy 553 ttgggaatgcagctctcaagaggtttctcoattggtgtacagatggcatagtgaagggcc 612  
|||||  
Db 361 TTTGGGAATCGAGTGTCTCAAGAGGTTCTCCATGGTGGTACAGATGGCATAGTGAAGGCC 420  
|||||

Qy 613 ctgaatgtgaaccacatggcagagcctcactcagcctggcagcccaaatatcatctca 672  
|||||  
Db 421 CTGAATGTGGAACAGATGCACAGGCTCCACCTGCAGGCTGCACCCCAATATCATCTCA 480  
|||||

Qy 673 cagctctga 681  
|||||  
Db 481 CAGCTCTGA 489  
|||||

RESULT 11  
A99105 LOCUS  
DEFINITION Sequence 9 from Patent WO9909054.  
ACCESSION A99105  
VERSION A99105.1 GI:6782057

KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 604)  
AUTHORS Falmagne,P., Wattiez,R., Bernard,A., Hermans,C. and Knoops,B.  
TITLE Peroxisome-associated polypeptide, nucleotide sequence encoding  
said polypeptide and their uses in the diagnosis and/or the  
treatment of lung injuries and diseases, and of oxidative  
stress-related disorders  
JOURNAL PATENT: WO 9909054-A 9 25-FEB-1999;  
UNIV MONS HAINAUT (BE) ; FALMAGNE PAUL (BE) ; WATTIEZ RUDDY (BE) ;  
BERNARD ALFRED (BE) ; HERMANS CEDRIC (BE) ; KNOOPS BERNARD (BE) ;  
UNIV LOUVAIN (BE)  
FEATURES  
source 1..604  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
CDS 161..517  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAB70594.1"  
/db\_xref="GI:6782058"  
/translation="MAPIKVGDAIPAVEVEFEGEPGNKNVLAELFKGKGVLFVGPVGF  
TPGCKSVRLPADPTGAFGKETDLLDDSLVSIFGNRRLRKFSMVVDGIVKALNVEPD  
GTGLTCSLAPNIISOL"

BASE COUNT 117 a 160 c 200 g 127 t  
ORIGIN

Query Match 57.4%; Score 462; DB 9; Length 604;  
Best Local Similarity 82.1%; Pred. No. 1.4e-106;  
Matches 604; Conservative 0; Mismatches 0; Indels 132; Gaps 1;

Qy 33 gggtagtgactagctggcgtgtgtgcgcctgagacgtcagcggtctatactcgtcg 92  
|||||  
Db 1 GGGTATGGAGCTAGTGGCGCTGTGCGCCTGAGAGCCTCAGCGGCTATATCTGCTCGG 60  
|||||

Qy 93 tgggcccgcggtcagctcgcgcagcgcagcaagcgtgcagtgagagagagtgagc 152  
|||||  
Db 61 TGGGCCCGGGTCACTAGTCTGGCGCAGCGCAGCAAGAGCGGTGACGTGAAGGAGAGTGCGC 120  
|||||

Qy 153 gtctggcggggtccgcagctttcagcagagcgcgtgcagcccatgcccacatcaaggtgg 212  
|||||  
Db 121 GTCTGGCGGGTCCCGAGCTTTTCAGCAGAGCGCTGCAGCCATGGCCCCAATCAAGGTGGG 180  
|||||

Qy 213 agatgccatccacagcagtgagggtgtttgaaggggagccaggggaacaggtgaacctggc 272  
|||||  
Db 181 AGATGCCATCCAGCAGTGGAGGTGTTTGAAGGGAGCGCAGGGAACAAGGTGAACCTGGC 240  
|||||

Qy 273 agagctgttcaaggccaagaggtgtgtgttttggagttccttgggccttcacccctgg 332  
|||||  
Db 241 AGAGCTGTTTCAAGGGCAAGAAGGGTGTGTGTTGGAGTCTCTGGGGCTTTCACCCCTGG 300  
|||||

Qy 333 atgttccaagacacacctggcagggtttgtgagcaggtgaggtctgaaggccaagg 392  
|||||  
Db 301 ATGTTC----- 306  
|||||

Qy 393 agtccaggtgtgtggtctgtctgtatgttaagtatgcttctgtgactggcgagtgggcg 452  
|||||  
Db 307 ----- 306  
|||||

Qy 453 agccccacaaggcggaaggttcggtcgtcgtcgtatccactgggcttttggaa 512  
|||||  
Db 307 -----CAAGGTTCCGCTCTGCTGATCCCATCCATGGGGCTTTTGGGAA 348  
|||||

Qy 513 ggagacagacttattactagatgattcgtgtgtgttccatctttgggaatcagcgtctcaa 572  
|||||  
Db 349 GGAGACAGACTTATTACTAGATGATTCGCTGGTGTCCATCTTTGGGAATCAGCTCTCAA 408  
|||||

Qy 573 gaggttctccatgtgtgttacaggatggcatagtgaggccctcctgaatgtggaaccagatgg 632  
|||||



RESULT 14







OM of: US-09-486-167A-2 to: Issued\_Patents\_NA:\* out\_format : pfs  
 Date: Oct 22, 2001 10:42 AM  
 About: Results were produced by the GenCore software, version 4.5,  
 Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

-MODEL=frame2.p2n.model -DEV=xlp  
 -O=/cgn2\_1/USPTO.spool/US09486167/runat\_22102001\_064207\_25045/app\_query.fasta\_1.222  
 -DB=Issued\_Patents\_NA -OFFWT=fastap -SUFFIX=rni -GAPOP=12.000  
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000  
 -QAPOP=4.500 -QAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
 -FCGPOP=6.000 -FCAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62  
 -TRANS=human40.cgi -LIST=45 -DOCLIGN=200 -THR\_SCORE=pct  
 -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
 -NORM-ext -MINLEN=0 -MAXLEN=2000000000  
 -USER=US09486167@cgn1\_1\_58 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLPXP  
 -WAIT -THREADS=1

## Search information block:

Query: US-09-486-167A-2  
 Query length: 162  
 Database: Issued\_Patents\_NA:\*  
 Database sequences: 324599  
 Database length: 94655562  
 Search time (sec): 57.360000

## score\_list:

Sequence	Strid	Orig	zScore	Escore	Len	Documentation
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-959-004-2 +	830.00	1877.26	6.0e-97	993	1	993
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-467-265-1 +	96.00	194.46	0.0032	918	1	918
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-467-265-1 +	96.00	194.46	0.0032	918	1	918
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-299-162A-1 +	93.50	193.48	0.0037	600	1	600
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-431-080-21 +	88.00	177.55	0.0284	807	1	807
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-431-080-21 +	88.00	177.55	0.0284	807	1	807
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-938-534-21 +	88.00	177.55	0.0284	807	1	807
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-938-534-21 +	88.00	177.55	0.0284	807	1	807
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-411-578-41 +	83.50	169.28	0.0820	672	1	672
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-036-987A-1 +	80.00	107.74	0.1948	80161	1	80161
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-602-262-3 +	77.50	156.78	0.4069	600	1	600
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-602-262-3 +	77.50	156.78	0.4069	600	1	600
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-004-716-3 +	77.50	156.78	0.4069	600	1	600
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-004-716-3 +	77.50	156.78	0.4069	600	1	600
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-602-262-3 +	77.50	154.48	0.5467	737	1	737
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-602-262-3 +	77.50	154.48	0.5467	737	1	737
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-004-716-3 +	77.50	154.48	0.5467	737	1	737
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-004-716-3 +	77.50	154.48	0.5467	737	1	737
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-280-443-1 +	77.50	129.83	12.91	6671	1	6671
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-457-459-1 +	77.50	129.83	12.91	6671	1	6671
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-555-678-1 +	77.50	129.83	12.91	6671	1	6671
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-320-878-23 +	73.50	136.00	5.85	1693	1	1693
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-939-218A-1 +	73.50	128.93	14.48	3183	1	3183
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-940-661A-1 +	73.50	128.90	14.54	3192	1	3192
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-09-083-485-1 +	73.50	128.90	14.54	3192	1	3192
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-151-189-1 +	72.50	127.56	17.26	2931	1	2931
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-125-468-1 +	72.50	101.53	486.42	30001	1	30001
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-474-933-1 +	72.50	101.53	486.42	30001	1	30001
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-09-090-567-1 +	72.00	137.58	4.78	1081	1	1081
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-415-751-32 +	71.00	135.58	6.18	1053	1	1053
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-415-751-32 +	71.00	135.58	6.18	1053	1	1053
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-258-349-1 +	71.00	122.58	32.70	3363	1	3363
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-320-878-19 +	71.00	95.30	1.4e+03	38506	1	38506
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-594-031-65 +	70.50	147.35	1.36	332	1	332
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-5171680-4 +	70.00	141.03	3.07	527	1	527
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-883-036A-1 +	69.50	129.58	13.32	1323	1	1323
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-329-633A-1 +	69.50	126.14	20.71	1799	1	1799
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-987-123-1 +	69.00	124.04	27.11	1959	1	1959
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-987-123-3 +	69.00	124.04	27.11	1959	1	1959

/cgn2\_6/ptodata/2/ina/5A.COMB.seq:US-07-961-522-1 + 69.00 120.47 42.88 2696  
 /cgn2\_6/ptodata/2/ina/5A.COMB.seq:US-08-217-438-1 + 69.00 120.47 42.88 2696  
 /cgn2\_6/ptodata/2/ina/5A.COMB.seq:US-08-321-978-1 + 69.00 120.47 42.88 2696  
 /cgn2\_6/ptodata/2/ina/5B.COMB.seq:US-08-710-584-1 + 69.00 120.47 42.88 2696  
 /cgn2\_6/ptodata/2/ina/6A.COMB.seq:US-08-961-083-43 + 68.50 139.79 3.60 433

seq\_name: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:US-08-959-004-2

seq\_documentation\_block:  
 ; Sequence 2, Application US/08959004  
 ; Patent No. 6197543  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hillman, Jennifer L.  
 ; APPLICANT: Yue, Henry  
 ; APPLICANT: Corley, Neil C.  
 ; APPLICANT: Lal, Preeti  
 ; APPLICANT: Shah, Purvi  
 ; APPLICANT: Kasez, Matthew  
 ; TITLE OF INVENTION: HUMAN VESICLE MEMBRANE PROTEIN-LIKE  
 ; NUMBER OF SEQUENCES: 11  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/959,004  
 ; FILING DATE: Herewith  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Billings, Lucy J.  
 ; REGISTRATION NUMBER: 36,749  
 ; REFERENCE/DOCKET NUMBER: PF-0414 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 650-855-0555  
 ; TELEFAX: 650-845-4166  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 993 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: BRAIT01  
 ; CLONE: 743725  
 ; US-08-959-004-2

alignment\_scores:  
 Quality: 830.00 Length: 162  
 Ratio: 5.123 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-486-167A-2 x US-08-959-004-2 ..  
 Align seg 1/1 to: US-08-959-004-2 from: 1 to: 993  
 1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 385 ATGCCCCCAATCAAGTGGGAGATGCCATCCACGACGTGGGTGTTGA 434

```
17 uGlyluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
|||||
435 AGGGAGCCAGGAAACAGTGAACCTGGCAGAGCTGTTCAAGGCAAGA 484
|||||
34 ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
|||||
485 AGGGTGTGCTGTTTGGAGTTCCTGGGGCTTCACCCCTGGATGTTCCAAG 534
|||||
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
|||||
535 ACACACCTGCCAGGTTGTGGAGCAGGCTGAGGCTCTGAAGCCCAAGG 584
|||||
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
|||||
585 AGTCCAGGTGGTGGCTGCTGAGTGTAAATGATGCCITTTGGACTGGCG 634
|||||
84 LuTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
|||||
635 AGTGGGGCGGAGCCACACAGGCGAAGCAAGGTTCCGGCTCTGGCTGAT 684
|||||
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117
|||||
685 CCCATGGGGCTTTGGAGAGACAGACTTATTACTAGATGATCGCT 734
|||||
117 uValSerLePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
|||||
735 GGTGTCCATCTTTGGAAATCGACGCTCTCAAGAGGTTCTCCATGGTGTAC 784
|||||
134 InAspGlyLeValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
|||||
785 AGATGGGATGATGAAGGCGGCGAATGGAACAGATGGCAGGCGCTC 834
|||||
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162
|||||
835 ACCTGCAGCGTGGCACCAATATCATCTCACAGCTC 870
|||||
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-467-265-1
seq_documentation_block:
; Sequence 1, Application US/08467265
; Patent No. 5985612
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Gentz, Reiner
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: NATURAL KILLER CELL ENHANCED FACTOR-C
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CRECHI,
; ADDRESSEE: STUART & OLSTEIN
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,265
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Feiraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-456
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
```

```
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 918 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31..843
; US-08-467-265-1

alignment_scores:
Quality: 96.00 Length: 153
Ratio: 1.067 Gaps: 8
Percent Similarity: 58.824 Percent Identity: 28.105

alignment_block:
US-09-486-167A-2 x US-08-467-265-1 ..
Align seg 1/1 to: US-08-467-265-1 from: 1 to: 918

2 AlaProIleLysValGlyAspAlaIleProAlaValGluValPheGluG1 18
|||||
283 GCGCCCTACTGGGAAGAACACGCT.....GTGATCGATGG 317
|||||
18 yGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysLysG 35
|||||
318 AGAA...TTTAAGGAGCTGAAGTTAAGTAT...TATCGTGGAAATACT 361
|||||
35 lyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLysThr 51
|||||
362 TGGTTTCTCTCTTACCCACTTGATTTTCACATTTGTGT...CCAAC 408
|||||
52 HisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysGlyVa 58
|||||
409 GAATTTATCGCTTTTGGCAGACAGACTTGAAGAAATTCAGATCTATAATAC 458
|||||
68 lGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyGlu 85
|||||
459 TGAAGTGGTAGCATGC...TCGTGTGATTCACAGTTTACCCATTTGGCCT 505
|||||
85 rpGlyArgAlaHisLysAlaGluGly.....LysValArg 96
|||
506 GGATTAATACCCCTCGAAGACAAAGGAGGACTTGGGCCAATAAGGATCCA 555
|||||
97 LeuLeuAlaAspProThrGlyAlaPheGlyLysGluThrAspLeuLeu 113
|||||
556 CTCTCTTTCAGATTTGACCCATCAGATCTCAAGAGGACTATGGTGTATACCT 605
|||||
113 uAspAspSerLeuValSerIlePheGlyAsnArgArgLeuLysArgPheS 130
|||||
606 AGAGGACTCA.....GGCCACACTCTTAGAGGCTCTCT 637
|||||
130 erMetValValGlnAspGlyLysValLys.....AlaLeuAsnValGlu 144
|||||
638 TCATTATTGATGACAAAGGAATCTCTAAGACAAATTAATCTCTGAATGATCTT 687
|||||
145 ProAspGly 147
|||
688 CCTGTGGGT 696

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-467-265-1
seq_documentation_block:
; Sequence 1, Application US/08467265
; Patent No. 6255079
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Gentz, Reiner
; APPLICANT: Rosen, Craig A.
```

```

; TITLE OF INVENTION: NATURAL KILLER CELL ENHANCED FACTOR-C
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,265
; FILING DATE: 06-Jun-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-456
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 918 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31..843
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-467-265-1

```

```

alignment_scores:
  Quality: 96.00      Length: 153
  Ratio: 1.067       Gaps: 8
  Percent Similarity: 58.824   Percent Identity: 28.105

```

```

alignment_block:
US-09-486-167a-2 x US-08-467-265-1 ..

Align seg 1/1 to: US-08-467-265-1 from: 1 to: 918

2 AlaProileLysValGlyAspAlaIleProAlaValGluValPheGluGl 18
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
283 GCSCCTACTGGGAAGGACAGT.....GTGATCGATGG 317

18 yGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysG 35
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
318 AGAA...TTAAGGAGCTGAAGTACTGAT..TATCGTGGGAATACT 361

35 lyValLeuPheGlyValProGlyAlaPheProGlyCysSerLysThr 51
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
362 TGGTTTCTCTCTACCCACTGATTTTCACATTTGTGT...CCAAC 408

52 HisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysGlyVa 68
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
409 GAAATTCGCTTTGGCGACAGACTTGAAGAATTCAGATCTATAAATAC 458

68 lGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyGlu 85
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
459 TGAAGTGGTAGCATGC...TCTGTTGATTACAGATTACCCCATTTGGCCT 505

85 rpGlyArgAlaHisLysAlaGluGly.....LysValArg 96
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
506 GGATTATTACCCCTCGAAGACAGGAGGACTTGGGCCAATAAGGATTCCA 555

```

```

97 LeuLeuAlaAspProThrGlyAlaPheGlyLysGluThrAspLeuLeu 113
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
556 CTTCTTTTCAGATTGACCCATCAGATCTCAAGGACTATGGTGTATACCT 605

113 uAspSerLeuValSerIlePheGlyAsnArgArgLeuLysArgPheS 130
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
606 AGAGGACTCA.....GGCCACACTCTTAGAGGTCTCT 637

130 erMetValValGlnAspGlyIleValLys.....AlaLeuAsnValGlu 144
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
638 TCATTATTGATGACAAAGGATCTCTAAGACAAATTACTCTGAATGATCTT 687

145 ProAspGly 147
||| |||
688 CCTGTGGGT 696

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-299-162A-1

seq_documentation_block:
; Sequence 1, Application US/08299162A
; Patent No. 5610286
; GENERAL INFORMATION:
; APPLICANT: Shau, Hungyi
; APPLICANT: Golub, Sidney H
; TITLE OF INVENTION: Natural Killer Cell Enhancing Factor
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David J. Oldenkamp
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/299,162A
; FILING DATE: August 31, 1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Oldenkamp, David J
; REFERENCE/DOCKET NUMBER: 104-280
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310-788-5000
; TELEFAX: 310-277-1297
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; POSITION IN GENOME:
; UNITS: bp
; NAME/KEY: CDS
; LOCATION: 1..532
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..600
US-08-299-162A-1

```

```

alignment_scores:
  Quality: 93.50      Length: 148

```

```

; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-431-080-21

alignment_scores:
    Quality: 88.00      Length: 133
    Ratio: 1.086      Gaps: 6
    Percent Similarity: 60.902      Percent Identity: 25.564

alignment_block:
    US-09-486-167A-2 x US-08-431-080-21/rev ..

    Align seg 1/1 to reverse of: US-08-431-080-21 from: 1 to: 807

        4 llelysalGlyAspAlaIleProAlaValGluValPheGluGlyGluPr 20
          ::::||||| ||||| ::::||||| ::::||||| ::::|||||
        567 TTAGATAGGCGATCCTATTCCCTGATTGAGTCTTTTA...AATGAAGA 521

        20 oGlyAsnlysalAsnLeuAlaGluLeuPheLysGlyLysGlyValL 37
          ::::||||| ||||| ::::||||| ::::||||| ::::|||||
        520 TAATGACTCTATCTCTTTGAAGAATAATCACCGAAATAACAGAGTGTGG 471

        37 euPheGlyVal...ProGlyAlaPheThrProGlyCysSerLysThrHis 52
          ::::||||| ||||| ||||| ||||| ||||| ||||| ::::
        470 TGTTTTGTGTATCCAGGCCAGCAGCCCTGGTTGCTACTAGA...CAG 424

        53 LeuProGlyPheValGluGlnAlaGluLeuLysAlaLysGlyValG 69
          :||| ||||| :::: ||||| :::: ||||| :::: ||||| ::::
        423 GCCTGTGGATTTCGTGACAATTACGAGGAAGTCAAG.....AA 386

        69 nValValAlaCysLeuSerValasnAspAlaPheValThrGlyGluTrpG 86
          ::::||||| ::::||||| ::::||||| ::::||||| ::::
        385 ATATGCTGCTCTCTTGGACTGAGTGCAGATTCGTGACATCCCAAGAAA 336

        86 lyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAspProThr 102
          ||||| :::: ::::||||| ::::||||| ::::||||| ::::
        335 AGTTTCAGAGTAACAAAATTTGCCATATCATTTACTAAGCGATCCAAAG 286

        103 GlyAlaPhe.....GlyLysGluThrAspLeu.....Le 112
          ||||| :::: ||||| :::: ||||| :::: ||||| ::::
        285 AGAGAGTTTATTGGTTGCTAGGAGCCAAAACCGCCATTTCTGGTTC 236

        112 uLeuAspAspSerLeuValSerIlePheGlyAsnArgLeuLysArg 128
          :::: ||||| :::: ||||| :::: ||||| :::: ||||| ::::
        235 TATTATGTCGATTTTCATTTTGTGGTGGGAAGTTAAATTCATAAAGA 187

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-431-080-31

seq_documentation_block:
; Sequence 31, Application US/08431080
; Patent No. 5698686
; GENERAL INFORMATION:
; APPLICANT: Gottschling, Daniel E.
; APPLICANT: Singer, Miriam S.
; TITLE OF INVENTION: Telomerase Compositions and Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433

```

```
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-431-080-21

alignment_scores:
    Quality:      88.00          Length:      133
    Ratio:        1.086         Gaps:           6
    Percent Similarity: 60.902   Percent Identity: 25.564

alignment_block:
US-09-486-167A-2 x US-08-431-080-21/rev ..

Align seg 1/1 to reverse of: US-08-431-080-21 from: 1 to: 807

4 llelysalGlyAspAlalProAlavalGluValPheGluGlyGluPr 20
:::||||| ||||| :::::::::::::: ::||
567 TTAGATAGCGCATCTTATTCCTGATTGTAGTCTTTTA...AATGAAGA 521

20 oGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysGlyValL 37
:::||||| ::||| :::::::::::::: ::||
520 TAATGACTCTATCTCCTTTGAAGAATAACCCGAAATAAACAGATGTGG 471

37 euPheGlyVal...ProGlyAlaPheThrProGlyCysSerLysThrHis 52
::||| ||| ||| ||| ||||| ||||| ::|
470 TGTTTTGTTGTCATCCCAGGCACGACGCCCTGGTTGCTACTAGA...CAG 424

53 LeuProGlyPheValGluGlnAlaGluLeuLysAlaLysGlyValGI 69
||||| ::| ::| ::||| ::|
423 GCCTGTGGATTTCGTGACAATTACCAGGAACCTCAAG.....AA 386

69 nValValAlaCysLeuSerValasnAspAlaPheValThrclyGluTrpG 86
:::|||| :::::::::::::: | ||||| ::|
385 ATATGCTGCTCTCTTGAGCTAGTGCGATTCGTGACATCCCAAGAAA 336

86 lYArgAlaHiSLysAlaGluGlyLysValArgLeuLeuAlaAspProthr 102
||| ::| ::| ::||| ::| |||||
335 AGTTTCAGAGTAAACAAAAATTGGCATATCATTTACTAAGCGATCCAAG 286

103 GlyAlaPhe.....GlyLysGluThrAspLeu.....Le 112
|||
285 AGAGAGTTTATTGGTTGCTAGGAGCCAAAAAACGCCACTTCTGGTTC 236

112 uLeuAspAspSerLeuValSerllePheGlyAsnArgLeuLysarg 128
::|
235 TATTATGTCGATTTTCATTTTTTTGTTGATGGGAAGTTAAAAATCAAAGA 187

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-431-080-31

seq_documentation_block:
; Sequence 31, Application US/08431080
; Patent No. 5698686
; GENERAL INFORMATION:
; APPLICANT: Gottschling, Daniel E.
; APPLICANT: Singer, Miriam S.
; TITLE OF INVENTION: Telomerase Compositions and Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
```



seq\_name: /cqn2\_6/ptodata/2/ina/6B\_COMB.seq:US-09-411-578-41

alignment_scores:		
Quality:	83.50	Length: 156
Ratio:	1.044	Gaps: 7
Percent Similarity:	51.282	Percent Identity: 25.000

alignment\_block:  
US-09-486-167A-2 x US-09-411-578-41  
Align seg 1/1 to: US-09-411-578-41

```
3  ProIleIysValGlyAspAlaIleProAlaValGluValPheGluGlyG1 19
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
4  CGCTTGAACCTGGAGATTCTTCCAGACACTTCCAGCG...GAGCGCT 50
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
19 uProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysGlyV 36
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
51 GGCGCGGAGCACTTCGCTTGCACAGTACTTGGGGGACACTGGGGAG 100
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
36 alLeuPheGlyValProGlyAlaPheThrProGlyCysSerLysThrHis 52
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
101 TGATGTTACGCCACCGAAGCACTTCCATCCCGTTTGCACA...ACGGAG 147
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
53 LeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaGlyGlyValG1 69
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
148 CTCGCCAGCCGCTGAAGCTCCAGGCTCTCTTCACGAAGAAGACTGCAA 197
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
69 nValValAlaCysLeuSerValAsnAspAlaPheValThrGlyGluTIPG 86
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
198 ACTCGTTGGC...TTCCTCTGCACACCTCTGCAGCCACAGATGGG 244
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
86 lyArgAlaHisLysAla.....GluGlyLysVal.....Arg 96
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
245 CGAAGATATAATGCGCTATGCAGGCGGATCTGGGAACCTTGCCTTCC 294
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
97 LeuLeuAlaAspProThrGlyAlaPheGlyLysGluThrAspLeuLeu 113
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
295 CTCGTTGGGACCC.....AATAGGGAACCTG..... 309
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
113 uAspAspSerLeuValSerIlePheGlyAsnArgArgLeuLysArgPheS 130
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
310 .....AATAGGGAACCTG..... 321
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
130 erMetValValGlnAspGlyIleValLysAlaLeuAsnValGluProAsp 146
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
322 .....GCCGCGAGTTTGGGAATATGATCCTGCAGAAAAGGACAAAAAG 366
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
147 GlyThrGlyLeuThrCys 152
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
367 GGGCTGCTTGTGACTTGC 384
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```

seq\_name: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-09-036-987A-1

```
seq_documentation_block:
; Sequence 1, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merio, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; TITLE OF INVENTION: Production
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow AgroSciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-NAR-1998
; CLASSIFICATION: 435
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80161 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-036-987A-1

alignment_scores:
Quality: 80.00 Length: 175
Ratio: 0.930 Gaps: 12
Percent Similarity: 49.143 Percent Identity: 29.714

alignment_block:
US-09-486-167A-2 x US-09-036-987A-1 ..
Align seg 1/1 to: US-09-036-987A-1 from: 1 to: 80161

12 AlaValGluValPheGluGlyGluProGlyAsnLysValAsnLeuAlaG1 28
||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
65551 GCGCGGAGGGTTTCGAGGGTTATCTAGGTAAATGCC...ACGCGGGGAGG 65597
||||| ||||| ||||| ||||| ||||| ||||| |||||

28 uLeuPheLysGlyLysLysGlyValLeuPheGlyValProGly..... 42
65598 CGTGTTCGGGTGGGTTCGTTGCTATTCGTTTGGTTTCGAGGCTCTCGG 65647
||||| ||||| ||||| ||||| ||||| ||||| |||||

43 ..AlaPheThrProGlyCysSerLysThr.....HisLeuPro 54
||||| ||||| ||||| ||||| ||||| ||||| |||||

65648 TGACGGTGGATACGCGGTTCGTCGTCGTTGGTGGCGCTGCACCTGGCG 65697
||||| ||||| ||||| ||||| ||||| ||||| |||||

65698 GGT.....CAAGCAGTGGGTCTGTTGAGTGTGATCTGGC 65732
||||| ||||| ||||| ||||| ||||| ||||| |||||

65 ..AlaLysGlyValGlnValAlaCysLeuSerValAsnAspAlap 80
||||| ||||| ||||| ||||| ||||| ||||| |||||
65733 TCTTCGGGTGGTGTGACGGTGTGTCGCGGCGGATG.....T 65773
||||| ||||| ||||| ||||| ||||| ||||| |||||

80 heValThrGlyGluTrpGlyArgAlaHisLysAlaGluGlyLysValArg 96
||||| ||||| ||||| ||||| ||||| ||||| |||||
65774 TCGTGGAGTTTTCGCGTCAACGGGCTTGGCGCGGATGGCGGTGCAAG 65823
||||| ||||| ||||| ||||| ||||| ||||| |||||

97 .....LeuLeuAlaAspProThrGlyAlaPheGlyLysGluThrAs 110
||||| ||||| ||||| ||||| ||||| ||||| |||||
65824 TCGTTTCGGCGGCTGCGGATGGCACCGGT...TGGGAGAGGCGCGG 65870
||||| ||||| ||||| ||||| ||||| ||||| |||||

110 pLeuLeuLeuAspSerLeuValSerIlePheGlyAsnArgArgLeuL 127
||||| ||||| ||||| ||||| ||||| ||||| |||||
65871 CTGTGTGTG.....TTGG 65884
||||| ||||| ||||| ||||| ||||| ||||| |||||

127 ysArgPheSerMetValValGlnAspGly.....IleVal 138
||||| ||||| ||||| ||||| ||||| ||||| |||||
65885 AGCGGCTGCGGATGCGCGGCGCAACGGGACACGGGTCTGCGCGGTG 65934
||||| ||||| ||||| ||||| ||||| ||||| |||||

139 LysAlaLeuAsnValGluProAspGly.....ThrGlyLeuThrCys 153
||||| ||||| ||||| ||||| ||||| ||||| |||||
65935 CGGGGTACGGGTGAATCAGGATGGTCCGTCGAATGTTTGACG..... 65979
||||| ||||| ||||| ||||| ||||| ||||| |||||

153 rLeuAlaProAsnIleIleSerGln 161
||||| ||||| ||||| ||||| ||||| ||||| |||||
65980 ....GCGCGAATGGGCGCTCGCAG 66000
||||| ||||| ||||| ||||| ||||| ||||| |||||

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-602-262-4
```



```
582 AAAGCATTCATTGGTCAACCGCCCTTAATTTCAAACAACACAGCGTTGT 533
22 .....AsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
   : : : : : : : : : : : : : : : : : : : : : : : :
532 GAATGCGGATTTCAAGGAATTTCACTTTGTGAG...TTCAAGGAAAT 486
   : : : : : : : : : : : : : : : : : : : : : : : :
34 ysglyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
   : : : : : : : : : : : : : : : : : : : : : : : :
485 ATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 439
   : : : : : : : : : : : : : : : : : : : : : : : :
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG 67
   : : : : : : : : : : : : : : : : : : : : : : : :
438 ACAGAGATAATTCCTTTTCTGCTGATGATGCGGAGTTCAAAAAATTAG 389
   : : : : : : : : : : : : : : : : : : : : : : : :
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGly 84
   : : : : : : : : : : : : : : : : : : : : : : : :
388 TGTAGCTGTTATGGCATGC...TCAACTGATTCACATTTTTCACACCT 342
   : : : : : : : : : : : : : : : : : : : : : : : :
84 luTrpGlyArgAlaHisLysAlaGluGly.....LysVal 95
   : : : : : : : : : : : : : : : : : : : : : : : :
341 CATGGGTAATAACCGACCAAAATGGGTGGACTCGGTGAGATGAATATA 292
   : : : : : : : : : : : : : : : : : : : : : : : :
96 ArgLeuLeuAlaAspProThrGlyAlaPheGlyLysGluThrAspLeu 112
   : : : : : : : : : : : : : : : : : : : : : : : :
291 CCAATTCTTCTGATACCAATCATACATTTAGTAGGCGATATGGCGTGT 242
   : : : : : : : : : : : : : : : : : : : : : : : :
112 uLeuAspAspSerLeuValSerIlePheGlyAsnArgArgLeuLysArg 129
   : : : : : : : : : : : : : : : : : : : : : : : :
241 CAAGGAAGATGATGGCATTTGCTTACCGTGA.....T 210
   : : : : : : : : : : : : : : : : : : : : : : : :
129 heSerMetValValGlnAspGlyIleValLysAlaLeuAsnVal 143
   : : : : : : : : : : : : : : : : : : : : : : : :
209 TATTATCATTTGATCCAAAGGATTTTGGCAGAAATCACAATC 166
   : : : : : : : : : : : : : : : : : : : : : : : :
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:us-09-004-716-4
seq_documentation_block:
; Sequence 4, Application US/09004716
; Patent No. 6031077
; GENERAL INFORMATION:
; APPLICANT: Klimowski, Laura
; APPLICANT: Tripp, Cynthia A.
; TITLE OF INVENTION: PARASITIC HELMINTH LARVAL THIOLE SPECIFIC
; TITLE OF INVENTION: ANTIOXIDANT PROTEINS, NUCLEIC ACID MOLECULES, AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/004,716
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/602,262
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Conneil, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-45
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/863-9700
; TELEFAX: 303/863-0223
; INFORMATION FOR SEQ ID NO: 4:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-004-716-4
```

```
alignment_scores:
  Quality: 77.50      Length: 148
  Ratio: 0.994       Gaps: 6
  Percent Similarity: 52.703  Percent Identity: 22.973
```

```
alignment_block:
```

```
US-09-486-167A-2 x US-09-004-716-4 ..
```

```
Align seg 1/1 to: US-09-004-716-4 from: 1 to: 600
```

```
14 GluValPheGluGlyGluProGly..... 21
   : : : : : : : : : : : : : : : :
19 AAAGCATTCATTGGTCAACCGCCCTTAATTTCAAACAACACAGCGTTGT 68
   : : : : : : : : : : : : : : : :
22 .....AsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
   : : : : : : : : : : : : : : : :
69 GAATGCGGATTTCAAGGAATTTCACTTTGTGAG...TTCAAGGAAAT 115
   : : : : : : : : : : : : : : : :
34 ysglyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
   : : : : : : : : : : : : : : : :
116 ATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 162
   : : : : : : : : : : : : : : : :
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG 67
   : : : : : : : : : : : : : : : :
163 ACAGAGATAATTCCTTTTCTGATGATGCGGAGTTCAAAAAATTAGA 212
   : : : : : : : : : : : : : : : :
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGly 84
   : : : : : : : : : : : : : : : :
213 TGTAGCTGTTATGGCATGC...TCAACTGATTCACATTTTTCACACCT 259
   : : : : : : : : : : : : : : : :
84 luTrpGlyArgAlaHisLysAlaGluGly.....LysVal 95
   : : : : : : : : : : : : : : : :
260 CATGGTAAATACCGACCAAAATGGGTGGACTCGGTGAGATGAATATA 309
   : : : : : : : : : : : : : : : :
96 ArgLeuLeuAlaAspProThrGlyAlaPheGlyLysGluThrAspLeu 112
   : : : : : : : : : : : : : : : :
310 CCAATTCTTCTGATACCAATCATACATTTAGTAGGCGATATGGCGTGT 359
   : : : : : : : : : : : : : : : :
112 uLeuAspAspSerLeuValSerIlePheGlyAsnArgArgLeuLysArg 129
   : : : : : : : : : : : : : : : :
360 CAAGGAAGATGATGGCATTTGCTTACCGTGA.....T 391
   : : : : : : : : : : : : : : : :
129 heSerMetValValGlnAspGlyIleValLysAlaLeuAsnVal 143
   : : : : : : : : : : : : : : : :
392 TATTATCATTTGATCCAAAGGATTTTGGCAGAAATCACAATC 435
   : : : : : : : : : : : : : : : :
```

```
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:us-09-004-716-5
```

```
seq_documentation_block:
```

```
; Sequence 5, Application US/09004716
; Patent No. 6031077
; GENERAL INFORMATION:
; APPLICANT: Klimowski, Laura
; APPLICANT: Tripp, Cynthia A.
; TITLE OF INVENTION: PARASITIC HELMINTH LARVAL THIOLE SPECIFIC
; TITLE OF INVENTION: ANTIOXIDANT PROTEINS, NUCLEIC ACID MOLECULES, AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: U.S.A.
```

```
;;
;; ZIP: 80203
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/004,716
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/602,262
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Connell, Gary J.
;; REGISTRATION NUMBER: 32,020
;; REFERENCE/DOCKET NUMBER: 2618-45
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 303/863-9700
;; TELEFAX: 303/863-0223
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 600 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cdna
US-09-004-716-5

alignment_scores:
  Quality: 77.50      Length: 148
  Ratio: 0.994       Gaps: 6
  Percent Similarity: 52.703      Percent Identity: 22.973

alignment_block:
US-09-486-167A-2 x US-09-004-716-5/rev ..

Align seg 1/1 to reverse of: US-09-004-716-5 from: 1 to: 600

14 GluValPheGluGlyGluProGly.....21
   ::::: ||| :::::
582 AAAGCATTTCATTGGTCAACCGGCCCTAATTTCAAACAACACGCGTTGT 533
22 .....AsnLysValasnLeuAlaGluLeuPheLysGlyLysL 34
   ::::: ||| ::::: ||| ::::: ||| ::::: |||
532 GAATGGCGATTTTCAGGAAATTTCACTTTGTCAG...TTCAAAGGAAAT 486
34 ysglyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
485 ATGTGTCCTCTCTTTTATCCTCTCGATTCACATTTTCGTTTGC...CCA 439

51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysGI 67
   ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
438 ACAGAGATAATTCGTTTCTGATCGTATTCGCGAGTTCAAATAATAGA 389
67 yValGlnValAlaLysCysSerValAsnAspAlaPheValThrGlyC 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
388 TGTAGCTGTATGGCATGC...TCAACTGATTCACATTTTCACACCTTG 342
84 LuTrpGlyArgAlaHisLysAlaGluGly.....LysVal 95
   ||| ::::: ||| ::::: |||
341 CATGGTAATACCGACCGCAAAATGGGTGGACTCGGTGAGATGAATATA 292
96 ArgLeuLeuAlaAspProThrGlyAlaPheGlyLysGluThrAspLeu 112
   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
291 CCATTCCTGCTGATACCAATCATCAATAGTAGGCGATAGCGGTGCT 242
112 uLeuAspAspSerLeuValSerIlePheGlyAsnArgArgLeuLysArgP 129
   , | ::::: |||
241 CAAGGAAGATGATGCGATTCCTACCGTGA.....T 210
129 heSerMetValValGlnAspGlyIleValLysAlaLeuAsnVal 143
```

```
;;
;; ::::: ||| ::::: ||| ::::: ||| ::::: |||
209 TATTCATTCATTGATCCAAAAGGATTTTGGCACAATCAACAATC 166
seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-602-262-1

seq_documentation_block:
; Sequence 1, Application US/08602262
; Patent No. 5744593
; GENERAL INFORMATION:
; APPLICANT: Klimowski, Laura
; APPLICANT: Tripp, Cynthia A.
; TITLE OF INVENTION: PARASITIC HELMINTH LARVAL THIOL SPECIFIC
; TITLE OF INVENTION: ANTIOXIDANT PROTEINS, NUCLEIC ACID MOLECULES, AND USES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,262
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-45
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/863-9700
; TELEFAX: 303/863-0223
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 737 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; NAME/KEY: CDS
; LOCATION: 6..605
US-08-602-262-1

alignment_scores:
  Quality: 77.50      Length: 148
  Ratio: 0.994       Gaps: 6
  Percent Similarity: 52.703      Percent Identity: 22.973

alignment_block:
US-09-486-167A-2 x US-08-602-262-1 ..

Align seg 1/1 to: US-08-602-262-1 from: 1 to: 737

14 GluValPheGluGlyGluProGly.....21
   ::::: ||| ::::: ||| ::::: ||| ::::: |||
24 AAAGCATTTCATTGGTCAACCGGCCCTAATTTCAAACAACACGCGTTGT 73
22 .....AsnLysValasnLeuAlaGluLeuPheLysGlyLysL 34
   ::::: ||| ::::: ||| ::::: ||| ::::: |||
74 GAATGGCGATTTTCAGGAAATTTCACTTTGTCAG...TTCAAAGGAAAT 120
34 ysglyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 ATGTGTCCTCTCTTTTATCCTCTCGATTCACATTTTCGTTTGC...CCA 167
```

```

51  ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1  67
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
168 ACAGAGATAATGGCTTTTCTGCGTATTCCGGAGTTCAAAAATTAGA  217
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
67  yValGlnValValAlaCysLeuSerValAsnAsAlaPheValThrGlyG  84
   ||| ||| : : : ||||| ||| : : : : : : : : : : |||
218 TGTAGCTGTTATGGCATGC...TCAACTGATTCCATTTTTCACACCTG  264
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
84  LuTrpGlyArgAlaHisLysAlaGluGly.....LysVal  95
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
265 CATGGGTAAATACCGCGAAAAATGGTGACTCGTCAGATGAATATA  314
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
96  ArgLeuLeuAlaAspProThrGlyAlaPheGlyLysGluThrAspLeuLe  112
   ||| : : ||||| ||| : : : : : : : : : : : : : : : : : : :
315 CCATTTCTGCTGATACCAATCATACAATTAGTAGGGCATATGGCGTCT  364
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
112 uLeuAspAspSerLeuValSerIlePheGlyAsnArgLeuLysArgP  129
   | : : : : : : : : : : : : : : : : : : : : : : : : : :
365 CAAGGAAGATGATGGCATTTGCTTACCGTGA.....T  396
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
129 heSerMetValValGlnAspGlyIleValLysAlaLeuAsnVal  143
   : : : : : : : : : : : : : : : : : : : : : : : : : :
397 TATTATCATGATCCAAAAGGGATTTTGCACAAATCACAATC  440
   : : : : : : : : : : : : : : : : : : : : : : : : : :

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2001, 12:15:26 ; Search time 58.12 seconds  
(without alignments)  
2622.083 Million cell updates/sec

Title: US-09-486-167A-1  
Perfect score: 805  
Sequence: 1 gccagaggcggtggaag.....ttgtgttgcggaaaaaaa 805

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 324599 seqs, 94655562 residues

Word size : 0

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents\_NA:\*  
1: /cgn2.6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2.6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2.6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2.6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2.6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2.6/ptodata/2/ina/backfiles1.seq:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	798	99.1	993	4	US-08-959-004-2
2	17	2.1	37	1	Sequence 2, Appli
3	17	2.1	37	1	Sequence 19, Appl
4	17	2.1	37	3	US-08-383-743A-19
5	17	2.1	37	3	US-08-808-881-19
6	17	2.1	37	5	US-09-017-631-19
7	17	2.1	1280	4	PCT-US93-07116-19
8	17	2.1	1930	1	Sequence 19, Appl
9	16	2.0	44	1	US-09-060-756-4
10	16	2.0	44	1	US-07-982-112-1
11	16	2.0	270	1	Sequence 12, Appl
12	16	2.0	439	5	US-08-348-143-12
13	16	2.0	558	4	US-08-571-785-12
14	16	2.0	732	1	US-08-143-576-2
15	16	2.0	794	3	US-09-060-756-240
16	16	2.0	1084	2	Sequence 240, App
17	16	2.0	1084	2	Sequence 9, Appli
18	16	2.0	1094	2	Sequence 69, Appl
19	16	2.0	1094	2	Sequence 8, Appli
20	16	2.0	1094	2	Sequence 14, Appl
21	16	2.0	1094	2	Sequence 110, App
22	16	2.0	1094	2	Sequence 109, App
23	16	2.0	1094	2	Sequence 109, App
24	16	2.0	1094	2	Sequence 5, Appli
25	16	2.0	1094	2	Sequence 2, Appli
26	16	2.0	1094	2	Sequence 19, Appl
27	16	2.0	1094	2	Sequence 8, Appli
28	16	2.0	1094	2	Sequence 1, Appli
29	16	2.0	1094	2	Sequence 24, Appl

Sequence 11, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 23, Appli  
Sequence 20, Appli  
Sequence 20, Appli  
Sequence 38, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 8, Appli  
Sequence 1, Appli  
Sequence 8, Appli  
Sequence 1, Appli  
Sequence 8, Appli  
Sequence 8, Appli  
Sequence 8, Appli

## ALIGNMENTS

RESULT 1  
US-08-959-004-2  
: Sequence 2, Application US/08959004  
: Patent No. 6197543  
: GENERAL INFORMATION:  
: APPLICANT: Hillman, Jennifer L.  
: APPLICANT: Yue, Henry  
: APPLICANT: Corley, Neil C.  
: APPLICANT: Lal, Preeti  
: APPLICANT: Shah, Purvi  
: APPLICANT: Kaser, Matthew  
: TITLE OF INVENTION: HUMAN VESICLE MEMBRANE PROTEIN-LIKE  
: TITLE OF INVENTION: PROTEINS  
: NUMBER OF SEQUENCES: 11  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Incyte Pharmaceuticals, Inc.  
: STREET: 3174 Porter Drive  
: CITY: Palo Alto  
: STATE: CA  
: COUNTRY: USA  
: ZIP: 94304  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Diskette  
: COMPUTER: IBM Compatible  
: OPERATING SYSTEM: DOS  
: SOFTWARE: FASTSEQ for Windows Version 2.0  
: CURRENT APPLICATION DATA:  
: FILING DATE: Herewith  
: CLASSIFICATION: 514  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER:  
: FILING DATE:  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Billings, Lucy J.  
: REGISTRATION NUMBER: 36,749  
: REFERENCE/DOCKET NUMBER: PF-0414 US  
: TELEPHONE: 650-855-0555  
: TELEFAX: 650-845-4166  
: TELEX:  
: INFORMATION FOR SEQ ID NO: 2:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 993 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: IMMEDIATE SOURCE:  
: LIBRARY: BRAITUT01  
: CLONE: 743725

US-08-959-004-2

Query Match 99.1%; Score 798; DB 4; Length 993;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 798; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 gccaggagcgagtggaagtgccctggggcggtatggactagctgctgtgcgc 60
DB 193 gccaggagcgagtggaagtgccctggggcggtatggactagctgctgtgcgc 252
QY 61 ctgagacgctcagcgagctatctcgtggtggcgccggtcagctcgtgcgcagcg 120
DB 253 CTGAGACGCTCAGCGGCTATATCTCGGTGGGGCGGGTCACTGCTCGCGCAGCG 312
QY 121 gcagcaagacggtgcagtgagagagtggtggtgctggtggtggtggtggtggt 180
DB 313 GCACGACGAGCGGTGACAGTGAAGGAGTGGCGTCTGGCGGGTCCGACAGTTCACGACA 372
QY 181 gccgctcagcagcagtgcccaatcaagtggtggagatgccatcccgagtgagtggttt 240
DB 373 GCCGCTGCAGCAGTGGCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGT 432
QY 241 gaaggagagcgagggagcaagtgtaacctggcagagctgttcaagggcaagaggtgtg 300
DB 433 GAAGGGAGCCAGGGAACAAGTGAACCTGGCAGAGCTGTTCAAGGGCAAGAGGTGTG 492
QY 301 ctgttgagagtgctggggtccctccctggtggttccaaagacacacctgcaggggttt 360
DB 493 CTGTTTGGAGTTCCTGGGGCTTCACCCCTGGATGTTCCAAGACACACCTGCCAGGTTT 552
QY 361 gtggagcagctgaggtctctgaagcgaaggtgaggtggtggtggtggtggtggtggt 420
DB 553 GTGGAGCAGCTGAGGCTCTGAAGGCCAAGGAGTCCAGGTGGTGGCTGTCTGAGTGT 612
QY 421 aatgctccttctgactggcgagtggtggcgagcccaagggcggaagcaaggttcgg 480
DB 613 AATGATGCTTGTGACTGGCGAGTGGCGGAGCCACAGGCGGAAGGCAAGGTTCGG 672
QY 481 ctcttgctgactcaccacggggttgggagagagagagagagagagagagagagagag 540
DB 673 CTCCTGGCTGATCCCACTGGGGCTTGGGAAGGAGAGAGACTTATTACTAGATGATCG 732
QY 541 ctggtgtccatttttgggaatcgactctcaagaggttctccatggtggttacagagatggc 600
DB 733 CTGCTGTCCATTTTGGGAATCAGCTCTCAAGAGTTCCTCAATGGTGGTACAGAGTGGC 792
QY 601 atagtgagggcctgaatgtggaaccagatggcagagcctcaacctgcagcctggcacc 660
DB 793 ATAGTGAAGGCCCTGAATGTGAACCAAGATGGCAGAGGCTCACTGCAGCCTGGCACCC 852
QY 661 aatcatctcacagctctgagccctggcgagagattacttctccacccctccctatct 720
DB 853 AATATCATCTCAGAGCTCTGAGGCCCTGGGCGAGATTACTTCTCCACCCCTCCCTATCT 912
QY 721 cacttgccagccctgtgctgggcccctgcaattggaatgttgccagatttctgcaata 780
DB 913 CACCTGCCAGCCCTGTGCTGGGCCCTGCAATTTGGAATGTTGGCCAGATTTCTGCAATA 972
QY 781 aacactgtggttgccgg 798
DB 973 AACACTTGTGTTGGCG 990

```

RESULT 2

US-08-383-743A-19/C  
 Sequence 19, Application US/08383743A  
 Patent No. 5614184  
 GENERAL INFORMATION:  
 APPLICANT: Sytkowski, Arthur J.  
 APPLICANT: Grodberg, Jennifer  
 TITLE OF INVENTION: RECOMBINANT HUMAN ERYTHROPOIETIN WITH  
 TITLE OF INVENTION: ALTERED BIOLOGICAL ACTIVITY

NUMBER OF SEQUENCES: 22  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
 STREET: Two Militia Drive  
 CITY: Lexington  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02173  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/383,743A  
 FILING DATE: 08/383,743  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/113,080  
 FILING DATE: 26-AUG-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/920,810  
 FILING DATE: 28-JUL-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/920,810  
 FILING DATE: 28-JUL-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Granahan, Patricia  
 REGISTRATION NUMBER: 32,227  
 REFERENCE/DOCKET NUMBER: NEDH92-04A2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 861-6240  
 TELEFAX: (617) 861-9540  
 INFORMATION FOR SEQ ID NO: 19:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 37 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-383-743A-19

Query Match 2.1%; Score 17; DB 1; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 30;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 372 tgaggctctgaaggcca 388
DB 23 TGAGGCTCTGAAGGCCA 7

```

RESULT 3

US-08-808-881-19/C  
 Sequence 19, Application US/08808881A  
 Patent No. 6048971  
 GENERAL INFORMATION:  
 APPLICANT: Sytkowski, Arthur J.  
 APPLICANT: Grodberg, Jennifer  
 TITLE OF INVENTION: RECOMBINANT HUMAN ERYTHROPOIETIN WITH  
 TITLE OF INVENTION: ALTERED BIOLOGICAL ACTIVITY  
 FILE REFERENCE: NEDH92-04A2  
 CURRENT APPLICATION NUMBER: US/08/808,881A  
 CURRENT FILING DATE: 1997-02-28  
 EARLIER APPLICATION NUMBER: US 08/383,743  
 EARLIER FILING DATE: 1995-02-02  
 EARLIER APPLICATION NUMBER: US 08/113,080  
 EARLIER FILING DATE: 1993-08-26  
 EARLIER APPLICATION NUMBER: US 07/920,810  
 EARLIER FILING DATE: 1992-07-28  
 NUMBER OF SEQ ID NOS: 22  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 19

; LENGTH: 37  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic oligonucleotide encoding protein mutant.  
US-08-808-881-19

Query Match 2.1%; Score 17; DB 3; Length 37;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 372 tgaggctctgaagccca 388  
|||||  
Db 23 TGAGGCTCTGAAGGCCA 7

RESULT 4  
US-09-017-631-19/c  
; Sequence 19, Application US/09017631B  
; Patent No. 6153407  
; GENERAL INFORMATION:  
; APPLICANT: Sytkowski, Arthur J.  
; APPLICANT: Grodberg, Jennifer  
; TITLE OF INVENTION: ERYTHROPOIETIN WITH ALTERED BIOLOGICAL  
; FILE REFERENCE: NEDH92-04A22A  
; CURRENT APPLICATION NUMBER: US/09/017,631B  
; CURRENT FILING DATE: 1998-02-03  
; EARLIER APPLICATION NUMBER: US 08/808,881  
; EARLIER FILING DATE: 1997-02-28  
; EARLIER APPLICATION NUMBER: US 08/383,743  
; EARLIER FILING DATE: 1995-02-02  
; EARLIER APPLICATION NUMBER: US 08/113,080  
; EARLIER FILING DATE: 1993-08-26  
; EARLIER APPLICATION NUMBER: US 07/920,810  
; EARLIER FILING DATE: 1992-07-28  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 19  
; LENGTH: 37  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic oligonucleotide encoding protein mutant.  
US-09-017-631-19

Query Match 2.1%; Score 17; DB 3; Length 37;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 372 tgaggctctgaagccca 388  
|||||  
Db 23 TGAGGCTCTGAAGGCCA 7

RESULT 5  
PCT-US93-07116-19/c  
; Sequence 19, Application PC/TUS9307116  
; GENERAL INFORMATION:  
; APPLICANT: New England Deaconess Hospital  
; TITLE OF INVENTION: RECOMBINANT HUMAN ERYTHROPOIETIN WITH  
; TITLE OF INVENTION: ALTERED BIOLOGICAL ACTIVITY  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02173  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/07116  
; FILING DATE: 19930728  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/920,810  
; FILING DATE: 28-JUL-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Granahan, Patricia  
; REGISTRATION NUMBER: 32,227  
; REFERENCE/DOCKET NUMBER: NEDH92-04  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 861-6240  
; TELEFAX: (617) 861-9540  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 37 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; PCT-US93-07116-19

Query Match 2.1%; Score 17; DB 5; Length 37;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 372 tgaggctctgaagccca 388  
|||||  
Db 23 TGAGGCTCTGAAGGCCA 7

RESULT 6  
US-09-060-756-4  
; Sequence 4, Application US/09060756  
; Patent No. 6183957  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Stewart  
; APPLICANT: Buchrieser-Brosch, Roland  
; APPLICANT: Gordon, Stephen  
; APPLICANT: Billault, Alain  
; TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM  
; TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA  
; TITLE OF INVENTION: LIBRARY APPLICATION TO THE DETECTION OF MYCOBACTERIA  
; FILE REFERENCE: 3495-0169  
; CURRENT APPLICATION NUMBER: US/09/060,756  
; CURRENT FILING DATE: 1998-04-16  
; NUMBER OF SEQ ID NOS: 743  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 1280  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
US-09-060-756-4

Query Match 2.1%; Score 17; DB 4; Length 1280;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 88 gtcggtggggccgccc 104  
|||||  
Db 679 gtcggtggggccgccc 695

RESULT 7  
US-07-982-112-1/c  
; Sequence 1, Application US/07982112

; Patent No. 5346813  
; GENERAL INFORMATION:  
; APPLICANT: BODENMULLER, Heinz  
; APPLICANT: DESSAUER, Andreas  
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF A SMALL  
; TITLE OF INVENTION: CELL LUNG CARCINOMA  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nikolaus, Marmelstein, Murray & Oram  
; STREET: 655 Fifteenth Street N.W. Suite 330  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-5701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/982,112  
; FILING DATE: 19921125  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/617,102  
; FILING DATE: 21-NOV-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Chin, Monica F.  
; REGISTRATION NUMBER: 36,105  
; REFERENCE/DOCKET NUMBER: 564-1906  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)638-5000  
; TELEFAX: (202)638-4810  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1930 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 10..933  
; US-07-982-112-1

Query Match 2.18; Score 17; DB 1; Length 1930;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 333 atgttccaagacacacc 349  
|||||  
Db 1668 ATGTTCCAAGACACACC 1652

RESULT 8  
US-08-348-143-12/c  
; Sequence 12, Application US/08348143  
; Patent No. 5506205  
; GENERAL INFORMATION:  
; APPLICANT: TAJIMA, HISAO  
; APPLICANT: KITAGAWA, KOICHIRO  
; APPLICANT: OHNO, HIROYUKI  
; APPLICANT: UENO, TOSHIO  
; TITLE OF INVENTION: A No. 5506205el Polypeptide of Protein p140 and DNAs  
; TITLE OF INVENTION: encoding it  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
; STREET: 2100 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.

; ZIP: 20037-3202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/348,143  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 315806/1993  
; FILING DATE: 24-NOV-1993  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)293-7060  
; TELEFAX: (202)293-7860  
; TELEX: 6491103  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 44 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (synthetic)  
; US-08-348-143-12

Query Match 2.04; Score 16; DB 1; Length 44;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 575 ggtttcccatggtggt 590  
|||||  
Db 26 GGTTCCTCATGTGCT 11

RESULT 9  
US-08-571-785-12/c  
; Sequence 12, Application US/08571785  
; Patent No. 5804411  
; GENERAL INFORMATION:  
; APPLICANT: TAJIMA, HISAO  
; APPLICANT: KITAGAWA, KOICHIRO  
; APPLICANT: OHNO, HIROYUKI  
; APPLICANT: UENO, TOSHIO  
; TITLE OF INVENTION: A No. 5804411el Polypeptide of Protein p140  
; TITLE OF INVENTION: and DNAs encoding it  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
; STREET: 2100 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20037-3202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/571,785  
; FILING DATE: 13-DEC-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/348,143  
; FILING DATE: 23-NOV-1994  
; APPLICATION NUMBER: JP 315806/1993  
; FILING DATE: 24-NOV-1993  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)293-7060  
; TELEFAX: (202)293-7860  
; TELEX: 6491103

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 44 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (synthetic)  
US-08-571-785-12

Query Match 2.0%; Score 16; DB 1; Length 44;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 575 gttctccatggtggt 590  
|||||

Db 26 GGTTCTCCATGGTGT 11

## RESULT 10

US-08-143-576-2  
Sequence 2, Application US/08143576  
Patent No. 5643761

## GENERAL INFORMATION:

APPLICANT: Fisher, Paul B.

APPLICANT: Jiang, Hongping

TITLE OF INVENTION: METHOD FOR GENERATING A SUBSTRUCTED

TITLE OF INVENTION: CDNA LIBRARY AND USES OF THE GENERATED LIBRARY

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: John P. White, c/o Cooper & Dunham

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/143,576

FILING DATE: 25-OCT-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 43563/JPW/AKC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 977-9550

TELEFAX: (212) 664-0525

TELEX: 422523 COOP UI

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 270 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-143-576-2

Query Match 2.0%; Score 16; DB 1; Length 270;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 641 tcaactgcagctgpc 656  
|||||

Db 121 TCACCTGCAGCTGCC 136

## RESULT 11

US-09-060-756-240/c  
Sequence 240, Application US/09060756  
Patent No. 6183957

## GENERAL INFORMATION:

APPLICANT: Cole, Stewart

APPLICANT: Buchrieser-Brosch, Roland

APPLICANT: Gordon, Stephen

APPLICANT: Billault, Alain

TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM

TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA

TITLE OF INVENTION: LIBRARY APPLICATION TO THE DETECTION OF MYCOBACTERIA

FILE REFERENCE: 3495-0169

CURRENT APPLICATION NUMBER: US/09/060.756

CURRENT FILING DATE: 1998-04-16

NUMBER OF SEQ ID NOS: 743

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 240

LENGTH: 439

TYPE: DNA

ORGANISM: Mycobacterium tuberculosis

US-09-060-756-240

Query Match

Best Local Similarity 2.0%; Score 16; DB 4; Length 439;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 681 aggcctggggccagat 696  
|||||

Db 184 AGGCCCTGGGCCAGAT 169

## RESULT 12

PCT-US95-08295-9  
Sequence 9, Application PC/TUS9508295  
GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: BREAST SPECIFIC GENES AND PROTEINS

NUMBER OF SEQUENCES: 30

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/08295

FILING DATE: 30-JUN-1995

CLASSIFICATION:

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 492 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 2..490

PCT-US95-08295-9

Query Match 2.0%; Score 16; DB 5; Length 492;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 306 tggagttctctggggcc 321  
|||||

Db 184 TGGAGTTCTCTGGGGCC 199

## RESULT 13

US-09-446-504-69/c  
; Sequence 69, Application US/09446504  
; Patent No. 6218150  
; GENERAL INFORMATION:  
; APPLICANT: UEMORI, Takashi  
; APPLICANT: SATO, Yoshimi  
; APPLICANT: FUJITA, Tomoko  
; APPLICANT: MIYAKE, Kazuo  
; APPLICANT: MURAI, Hiroyuki  
; APPLICANT: ASADA, Kiyocho  
; APPLICANT: KATO, Ikunoshin  
; TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS  
; FILE REFERENCE: 1422-408PCT  
; CURRENT APPLICATION NUMBER: US/09/446,504  
; PRIOR FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: PCT/JP98/02845  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: JP 9-187496  
; PRIOR FILING DATE: 1997-06-26  
; PRIOR APPLICATION NUMBER: JP 9-320692  
; PRIOR FILING DATE: 1997-11-27  
; NUMBER OF SEQ ID NOS: 92  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 69  
; LENGTH: 558  
; TYPE: DNA  
; ORGANISM: Pyrococcus furiosus  
US-09-446-504-69

Query Match 2.0%; Score 16; DB 4; Length 558;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 569 tcaagaggttctccat 584  
Db 332 TCAAGAGGTTCTCCAT 317

RESULT 14  
US-08-145-006C-8  
; Sequence 8, Application US/08145006C  
; Patent No. 5656452  
; GENERAL INFORMATION:  
; APPLICANT: Rao, Anjana  
; APPLICANT: Hogan, Patrick Gerald  
; APPLICANT: McCaffrey, Patricia  
; APPLICANT: Jain, Jugnu  
; TITLE OF INVENTION: NF-ATp, A T LYMPHOCYTE  
; TITLE OF INVENTION: DNA-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/145,006C  
; FILING DATE: October 29, 1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/017,052  
; FILING DATE: February 11, 1993  
; APPLICATION NUMBER: 08/006,067  
; FILING DATE: January 15, 1993  
; ATTORNEY/AGENT INFORMATION:

; NAME: Fraser, Janis K.  
; REGISTRATION NUMBER: 34,819  
; REFERENCE/DOCKET NUMBER: 04590/007001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 732  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-145-006C-8

Query Match 2.0%; Score 16; DB 1; Length 732;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 306 tggagttctctggggcc 321  
Db 532 TGGAGTTCTCTGGGGCC 547

RESULT 15  
PCT-US94-00545-8  
; Sequence 8, Application PC/TUS9400545  
; GENERAL INFORMATION:  
; APPLICANT: Rao, Anjana  
; APPLICANT: Hogan, Patrick Gerald  
; APPLICANT: McCaffrey, Patricia  
; APPLICANT: Jain, Jugnu  
; TITLE OF INVENTION: NF-ATp, A T LYMPHOCYTE  
; TITLE OF INVENTION: DNA-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/00545  
; FILING DATE: 18-JAN-1994  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/145,006  
; FILING DATE: October 29, 1993  
; APPLICATION NUMBER: 08/017,052  
; FILING DATE: February 11, 1993  
; APPLICATION NUMBER: 08/006,067  
; FILING DATE: January 15, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fraser, Janis K.  
; REGISTRATION NUMBER: 34,819  
; REFERENCE/DOCKET NUMBER: 04590/007001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 732  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear

us-09-486-167a-1.oligo.rni

Mon Oct 22 14:46:47 2001

PCT-US94-00545-8

Query Match 2.0%; Score 16; DB 5; Length 732;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0; .

Qy 306 tggagttcctggggcc 321  
 |||||  
 Db 532 TGGAGTTCCTGGGGCC 547

Search completed: October 22, 2001, 13:44:27  
 Job time: 5341 sec



us-09-486-167a-1.oligo.rge

Mon Oct 22 14:46:45 2001

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: October 22, 2001, 12:13:50 ; Search time 1307.97 Seconds  
(without alignments)  
9519.743 Million cell updates/sec

Title: US-09-486-167A-1  
Perfect score: 805  
Sequence: 1 gccaggaggcggagtggaag.....ttgtgttgcggaaaaaaa 805

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1344157 seqs, 7733874588 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba1.\*  
2: gb\_ba2.\*  
3: gb\_ba3.\*  
4: gb\_in1.\*  
5: gb\_in2.\*  
6: gb\_in3.\*  
7: gb\_om.\*  
8: gb\_ov.\*  
9: gb\_pat1.\*  
10: gb\_pat2.\*  
11: gb\_ph.\*  
12: gb\_pl1.\*  
13: gb\_pl2.\*  
14: gb\_pl3.\*  
15: gb\_pl4.\*  
16: gb\_ba1.\*  
17: em\_ba2.\*  
18: em\_fun.\*  
19: em\_htgo\_hum.\*  
20: em\_htgo\_inv.\*  
21: em\_htgo\_rod.\*  
22: em\_htg\_hum1.\*  
23: em\_htg\_hum2.\*  
24: em\_htg\_hum3.\*  
25: em\_htg\_hum4.\*  
26: em\_htg\_hum5.\*  
27: em\_htg\_hum6.\*  
28: em\_htg\_hum7.\*  
29: em\_htg\_hum8.\*  
30: em\_htg\_inv1.\*  
31: em\_htg\_inv2.\*  
32: em\_htg\_other.\*  
33: em\_htg\_rod.\*  
34: em\_hum1.\*  
35: em\_hum2.\*  
36: em\_hum3.\*  
37: em\_hum4.\*  
38: em\_hum5.\*  
39: em\_hum6.\*  
40: em\_hum7.\*  
41: em\_in.\*  
42: em\_om.\*  
43: em\_or.\*

44: em\_ov.\*  
45: em\_pat.\*  
46: em\_ph.\*  
47: em\_pl.\*  
48: em\_ro.\*  
49: em\_sts.\*  
50: em\_sy.\*  
51: em\_un.\*  
52: em\_vl.\*  
53: gb\_sts1.\*  
54: gb\_sts2.\*  
55: gb\_sts3.\*  
56: gb\_sy.\*  
57: gb\_un.\*  
58: gb\_vl1.\*  
59: gb\_vl2.\*  
60: gb\_htg1.\*  
61: gb\_htg2.\*  
62: gb\_htg3.\*  
63: gb\_htg4.\*  
64: gb\_htg5.\*  
65: gb\_htg6.\*  
66: gb\_htg7.\*  
67: gb\_htg8.\*  
68: gb\_htg9.\*  
69: gb\_htg10.\*  
70: gb\_htg11.\*  
71: gb\_htg12.\*  
72: gb\_htg13.\*  
73: gb\_htg14.\*  
74: gb\_htg15.\*  
75: gb\_htg16.\*  
76: gb\_htg17.\*  
77: gb\_htg18.\*  
78: gb\_htg19.\*  
79: gb\_htg20.\*  
80: gb\_htg21.\*  
81: gb\_htg22.\*  
82: gb\_htg23.\*  
83: gb\_htg24.\*  
84: gb\_htg25.\*  
85: gb\_pr1.\*  
86: gb\_pr2.\*  
87: gb\_pr3.\*  
88: gb\_pr4.\*  
89: gb\_pr5.\*  
90: gb\_pr6.\*  
91: gb\_pr7.\*  
92: gb\_pr8.\*  
93: gb\_pr9.\*  
94: gb\_rol.\*  
95: gb\_rol.\*  
96: gb\_in4.\*  
97: gb\_pr10.\*  
98: em\_ba3.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	805	100.0	805	9	A99097
2	805	100.0	805	88	AF110731 Homo sapi
3	785	97.5	875	89	AF242525 Homo sapi
4	747	92.8	840	89	AF124993 Homo sapi
5	729	90.6	788	89	AF231705 Homo sapi
6	557	69.2	835	89	AF112212 Homo sapi
7	528	65.6	780	89	AF197952 Homo sapi
8	489	60.7	489	93	HSA249483



QY 241 gaaggagcagcaggaacaaggtgaacctggcagagctgttcaagggaaggaaggggtgtg 300  
Db 241 GAAGGGGAGCCAGGGAACAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGAAGGGTGTG 300  
QY 301 ctgtttggagtttctctggccctcacccttgatgttccaaacacacactgcagggttt 360  
Db 301 CTGTTTGGAGTTCCTCTGGGCTTCACCCCTGGATGTTTCAAGACACACACTGCCAGGGTTT 360  
QY 361 gtggagcagctgaggtctgaaggccaaggaggtccaggtggtggtggtggtggtggt 420  
Db 361 GTGAGCAGGCTGAGGCTCTGAAGGCCAAGGGAGTCCAGGTGGTGGCTGTCTGAGTGT 420  
QY 421 aatgatgcctttgtgactggcggcggcggcggcggcggcggcggcggcggcggcggc 480  
Db 421 AATGATGCTTTGTGACTGGCGGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480  
QY 481 ctctggtctatccactgggccccttgggaaggagacagacttattatagatgattcg 540  
Db 481 CTCCTGGCTGATCCACTTGGGCTTGGGAAGGAGAGACACTTATTACTAGATGATTCG 540  
QY 541 ctggtgtccatctttgggaatcgacgtctcaaggaggttccatggtggtggtggtggtg 600  
Db 541 CTGGTGTCCATCTTTGGGAATCGAGCTCTCAAGAGGTCTTCCATGTTGGTACAGGATGC 600  
QY 601 atagtgaaggccctgaatgtgaaccagatggcagagcctcactgcagcctggcacc 660  
Db 601 ATAGTGAAGGCCCTGAATGTGGAACAGATGCGACAGGCTCAGCTGCAGCCTGGCACCC 660  
QY 661 aatatactctcacagctctgagcctggcggcggcggcggcggcggcggcggcggcggc 720  
Db 661 AATATCATCTCACAGCTCTGAGGCGCTGGGCGCAGATTACTTCCCTCCACCCCTCCCTATCT 720  
QY 721 cactgcgcagcctgtgctgggcccctgcaattggaatgtggcagatttctgcaata 780  
Db 721 CACTGCCAGCCCTGCTGCTGGGCGCCTGCAATTGGAATGTTGGCAGATTCTTCTGCAATA 780  
QY 781 aacactgtggttcgggaagaaaaa 805  
Db 781 AACACTGTGTTGCGGGAAGAAAAA 805

RESULT 3  
AF242525 875 bp mRNA PRI 15-AUG-2000  
LOCUS Homo sapiens hypothetical protein SBB110 mRNA, complete cds.  
DEFINITION  
ACCESSION AF242525  
VERSION AF242525.1 GI:9802047  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 875)  
AUTHORS Zhang,W., Li,N., Wan,T. and Cao,X.  
TITLE Direct Submission  
JOURNAL Submitted (08-MAR-2000) Department of Immunology, Second Military  
Medical University & Shanghai Brilliance Biotechnology Institute,  
800 Xiangyin Rd., Shanghai 200433, P.R. China  
FEATURES  
source 1..875  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
36..680  
/codon\_start=1  
/product="hypothetical protein SBB110"  
/protein\_id="AAF99605.1"  
/db\_xref="GI:9802048"  
/translation="MGLAGVCAALRRSAGYILVGGAGGQSAARRCSEGEWASGGVR  
SFSRAAAMAPIKVGDAIPAVEFEGEPGNVLAELFKGKGVLFVGPAGFTPGCSK  
THLPGEVEAEALAKAQVQVAVCLSVNDAFTGEGRAHKAEGKRVLLADPTGAFGKE  
TDLLDLSLVSIFGNRLKRFMSVVDGIVKALNVDPDGTGLTCSLAPNIISQL"  
BASE COUNT 230 a 206 c 272 g 167 t  
ORIGIN

AF110731.1 GI:6103723  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 805)  
AUTHORS Wattiez,R., Hermans,C., Bernard,A., Lesur,O. and Falmagne,P.  
TITLE Human bronchoalveolar lavage fluid: two-dimensional gel  
electrophoresis, amino acid microsequencing and identification of  
major proteins  
JOURNAL Electrophoresis 20 (7), 1634-1645 (1999)  
MEDLINE 99351776  
PUBMED 10424490  
REFERENCE 2 (bases 1 to 805)  
AUTHORS Knoops,B., Clippel,A., Bogard,C., Arsalane,K., Wattiez,R.,  
Hermans,C., Duconseille,E., Falmagne,P. and Bernard,A.  
TITLE Cloning and characterization of AOB166, a novel mammalian  
antioxidant enzyme of the peroxidoreductase family  
JOURNAL J. Biol. Chem. 274 (43), 30451-30458 (1999)  
MEDLINE 99452929  
PUBMED 10521424  
REFERENCE 3 (bases 1 to 805)  
AUTHORS Knoops,R., Wattiez,R., Falmagne,P., Hermans,C. and Bernard,A.  
TITLE Direct Submission  
JOURNAL Submitted (04-DEC-1998) Department of Biology, Catholic University  
of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium  
FEATURES  
source 1..805  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="liq13; about 7cR from marker D11S913; between  
markers D11S1963 and D11S4407"  
37..681  
/note="AOB166; alkyl hydroperoxide reductase;  
thioredoxin peroxidase; represents Prxv peroxidoreductin  
subfamily in mammals"  
/codon\_start=1  
/product="antioxidant enzyme B166"  
/protein\_id="AAF03750.1"  
/db\_xref="GI:6103724"  
/translation="MGLAGVCAALRRSAGYILVGGAGGQSAARRCSEGEWASGGVR  
SFSRAAAMAPIKVGDAIPAVEFEGEPGNVLAELFKGKGVLFVGPAGFTPGCSK  
THLPGEVEAEALAKAQVQVAVCLSVNDAFTGEGRAHKAEGKRVLLADPTGAFGKE  
TDLLDLSLVSIFGNRLKRFMSVVDGIVKALNVDPDGTGLTCSLAPNIISQL"  
misc\_feature 193  
/note="probable alternative translation start site"  
BASE COUNT 163 a 200 c 276 g 166 t  
ORIGIN

Query Match 100.0%; Score 805; DB 88; Length 805;  
Best Local Similarity 100.0%; Pred. No. 0; Indels 0; Gaps 0;  
Matches 805; Conservative 0; Mismatches 0; Gaps 0;  
QY 1 gccagagcagcagtgaaagtcgctggcggcggtatggagctagctgctgctgctgctg 60  
Db 1 GCCAGAGCGGAGTGGAGTGGCGCGGTGGCGGCGGTATGGACTAGTGGCGGTGGCGG 60  
QY 61 ctgagacgctcagcgggctatatactcgtcgtggtggcgcggtcagtcggtcggcagcg 120  
Db 61 CTGAGACGCTCAGCGGGCTATATACTCGTGGTGGCGCGCGGTGAGTGGCGGCGGCG 120  
QY 121 gcagcaagcagcgtgagtgagagagagtgagcgtcgtcggcggtcccgagtttcagcaga 180  
Db 121 GCAGCAAGCAGCGTGGAGTGGAGAGTGGCGGCTGCTGGCGGGGTCCGAGTTTCAGCAGA 180  
QY 181 gccgctgcagcagcagtcgcccccaatcagtggtggagatgccatcccgagcagtggaagtg 240  
Db 181 GCCGCTGCAGCCATGCCCCCATCAATCAAGTGGGAGATGCATCCCGAGCAGGTGGAGTGT 240



541 ctgggtccatcttgggaatcgacgtctcaagaggttctccatggtggtacagatggc 600  
Db 574 CTGGGTCCATCTTTGGGAATCGACGCTCTCAAGAGGTTCTCCATGGTGTACAGATGGC 633  
QY 601 atagtgaaggccctgaatgtggaacacagatgacagagcctcaccctgagctggcacc 660  
Db 634 ATAGTGAAGGCCCTGAATGTGGAACAGATGGACAGAGCCCTCACTGAGGCTGGCACC 693  
QY 661 aatatacatcacagctctgagccctggccagattacttctctccacccctccatct 720  
Db 694 AATATCATCTACAGCTCTGAGGCCCTGGGCCAGATTACTTCTCCACCCCTCCCTATCT 753  
QY 721 cactgcccagccctgtgtggggccctcgaattggaatgttgccagattctgcaata 780  
Db 754 CACCTGCCAGCCGTGTGTGGGGCCCTGCAATTGGAATGTTGGCCAGATTCTTCAATA 813  
QY 781 aacactgtggtttgcg 798  
Db 814 AACACTGTGTTGGCG 831  
RESULT 5  
AF231705 AF231705 788 bp mRNA PRI 27-JUN-2000  
LOCUS Homo sapiens Alu co-repressor 1 (ACRI) mRNA, complete cds.  
DEFINITION AF231705  
ACCESSION AF231705  
VERSION AF231705.1 GI:8745393  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE Kropotov,A., Sedova,V., Ivanov,V., Sazeeva,N., Tomilin,A.,  
AUTHORS Krutillina,R., Oei,S.L., Griesenbeck,J., Buchlow,G. and Tomilin,N.  
TITLE A novel human DNA-binding protein with sequence similarity to a  
subfamily of reox proteins which is able to repress  
RNA-polymerase-III-driven transcription of the Alu-family  
retroposons in vitro  
JOURNAL Eur. J. Biochem. 260 (2), 336-346 (1999)  
MEDLINE 99195471  
PUBMED 10095767  
REFERENCE 2 (bases 1 to 788)  
AUTHORS Kropotov,A.V. and Tomilin,N.V.  
TITLE Direct Submission  
JOURNAL Submitted (04-FEB-2000) Chromosome Stability, Institute of Cytology  
RAS, Tikhoretskii Av.4, St.Petersburg 194064, Russia  
FEATURES  
source Location/Qualifiers  
1..788  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11q13"  
1..788  
/gene="ACRI"  
20..664  
/gene="ACRI"  
/codon\_start=1  
/product="Alu co-repressor 1"  
/protein\_id="AAF78899.1"  
/db\_xref="GI:8745394"  
/translation="MGLAGVCAALRRSAGYILVGGAGQGSAAAARCSGEWASGGVR  
SFSRAAAMAPIKVGDAPVAVFVEFEGPKNLAELFKGKGVLFQVPGAFPPGCSK  
THLPFGVEQALKAQGVVAVCLSVNDVFTGEWRAHKAEGKVRLLADPTGAFGKE  
TDLLDLSLVSIFCNRLKRFSSVMVDGIKALNVEPDGTGLTCLAPNIISQL"  
BASE COUNT 153 a 201 c 268 g 166 t  
ORIGIN  
Query Match 90.6%; Score 729; DB 89; Length 788;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 779; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 agtggcgtggggcggtatgggactagctggcgtgtgcccctgagacgctcagcgggc 78  
Db 2 AGTGGCGTGGGGCGGGTATGGGACTAGCTGGCGTGTGGCCCTGAGAGGCTCAGCGGC 61  
QY 79 tatatactctcgttggggcggtcagctcgtggcgacgagcagcagcagcagcagc 138  
Db 62 TATATACTCTCGTGGGGCGGGTCACTGCGGCGACGGCAGCAGCGGTGCAGT 121  
QY 139 gaagagagtgggcgcttggcggttcgcagtttccagcagagccgctgcagccatggcc 198  
Db 122 GAAGCAGAGTGGGCTCTGGGGGTCGCCAGTTTTCAGCAGAGCGCGCTGCAGCATGGCC 181  
QY 199 ccaatcaaggtgggagatgcatcccccagcagtggtttgaaggagcagcagcagc 258  
Db 182 CCAATCAAGGTGGGAGATGCCATCCAGCAGTGGAGGTGTTGAAGGGGAGCCAGGGAAC 241  
QY 259 aagtgaaacctggcagagctgttcaagggaagaggtgtgctgtttgagttcctggg 318  
Db 242 AAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGAGGTGTGCTGTTGGAGTTCTCTGG 301  
QY 319 gccttcacccctggatgttccaaagacacacctgcaggggtttgtggagcagcgtgaggtc 378  
Db 302 GCCITCACCCCTGGATGTTCCAAAGACACACCTGCACAGGGTTTGTGGAGCGTGAGGCT 361  
QY 379 ctgaaggccaaaggagtcaccaggtgtggcgtgtctgagtgtaatgatgcctttgtgact 438  
Db 362 CTGAAGGCCAAAGGAGTCCAGGTGGTGGCCTGTCTGAGTGTAAATGATGCCCTTTGTGACT 421  
QY 439 ggcagtgggcgccagcagcccaagggcaaggggttcggcctcctggctggtatcccaact 498  
Db 422 GCGAGTGGGGCCGAGCCCAAGCGGAGCAAGGTTCGGCTCCTGGCTGATCCCACT 481  
QY 499 gggcctttgggaaggagacagacttattactagatgattcgtggtgtccatcttggg 558  
Db 482 GGGGCTTTGGGAAGAGAGACAGACTTATTACTAGATGATTCCTGGTGTCCATCTTTGGG 541  
QY 559 aatcagctctcaagagtttccatggtgtacaggtggtacaggtgacaggtgacagcctgaat 618  
Db 542 AATCGACGCTCAAGAGGTCTCCATGTTGGTGGTACAGGATGGATAGTAGAGCCCTGAAT 601  
QY 619 gtggaaccagatggcacagcctcaccctgcagcgtggcaccacatatactcagcagctc 678  
Db 602 GTGGAAACAGATGGCACAGCGCTCACCCTGACGCTGGCAGCAATATATCTTCACAGCTC 661  
QY 679 tgaggccctggccagattacttctccacccctccctatctcactcagcagcctgtg 738  
Db 662 TGAGGCCCTGGCGCAGATTACTTCTCCACCCCTCCCTATCTACCTGCCCCAGCCCTGTG 721  
QY 739 ctgggcccctgcaattggaatgtggccagatttctgcaataaacacttgggttggg 798  
Db 722 CTGGGGCCCTGCANTTGAATGTTGCCAGATTCTTGCAATAAACACTTGTGGTTGCGG 781  
RESULT 6  
AF112212 AF112212 835 bp mRNA PRI 13-DEC-1999  
LOCUS Homo sapiens putative peroxisomal antioxidant enzyme mRNA, complete  
DEFINITION cds.  
ACCESSION AF112212  
VERSION AF112212.1 GI:6563211  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 835)  
AUTHORS Ren,S., Shi,J., Huang,C., Jiang,C., Li,Y., Zhou,J., Yu,Y., Xu,S.,  
Wang,Y., Fu,G., Chen,Z. and Han,Z.  
TITLE A novel gene expressed in human adrenal gland  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 835)  
AUTHORS Ren,S., Shi,J., Huang,C., Jiang,C., Li,Y., Zhou,J., Yu,Y., Xu,S.,  
Wang,Y., Fu,G., Chen,Z. and Han,Z.

TITLE Direct Submission  
JOURNAL Submitted (10-DEC-1998) Chinese National Human Genome Center at Shanghai, Guo Shoujing Rd. 351, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China

FEATURES  
Source Location/Qualifiers  
1. .835  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tissue\_type="adrenal gland"  
184. .672  
/note="PLP"  
/codon\_start=1  
/evidence=not\_experimental  
/product="putative peroxisomal antioxidant enzyme"  
/protein\_id="AAF17200.1"  
/db\_xref="GI:6563212"  
/translation="MAPIKVGDAIPAVEFEGEPGNKVNLAELFKGKGVLFQVPGAF  
TPGCKSTHLPVGEQNEALKAKGVQVVAFLSYNDAPVTGEMGRAHKAEGKRVLLADPT  
GARGKETDLLDDSLVSIFGNRLKRFMSVMVDGIVKALNVEPDGTGLTCLAPNIIS  
QL"

BASE COUNT 198 a 199 c 272 g 166 t

ORIGIN

Query Match 69.2%; Score 557; DB 89; Length 835;  
Best Local Similarity 100.0%; Pred. No. 5.6e-311;  
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 aagacgtgcagtgaggagagtgccgtctggcggggtcccgagtttcagcagagccgc 185  
|||||  
Db 117 AAGACGCTGCAGTGAAGGAGTGGCGCTGGCGGGTCCGAGTTTCAGCAGAGCCGC 176  
|||||

QY 186 tcagccatggcccaatcaagtgaggagatgccatcccgagtcgaggtgtttgaagg 245  
|||||

Db 177 TCAGCCATGGCCCCAATCAAGTGGAGATGCCATCCAGCAGTGGAGTGTTCGAAGG 236  
|||||

QY 246 ggagccagggaacaagtgaaactggcagagctgttcaaggccaagaggtgtctgtt 305  
|||||

Db 237 GGAGCCAGGGAACAAGTGAACCTGGCAGAGCTGTCAAGGGCAAGAGGTGTCTGT 296  
|||||

QY 306 tggagttcctggggccttaccctggatgtttcccaagacacacctgccaggtttgtga 365  
|||||

Db 297 TGAGTCTCTGGGGCTTCACCCCTGGATGTTCCAAAGACACACACCTGCCAGGTTTGTGA 356  
|||||

QY 366 gcaggtgaggtctgaagcccaaggagtgccagtggtgacctgtctgagtgtaatga 425  
|||||

Db 357 GCAGGCTGAGGCTCGAAGCCCAAGGAGTCCAGGTGTGGCTGTCTGAGTGTATGA 416  
|||||

QY 426 tgcccttgactggcagtgaggggccgagcccaaggcgaagggttcggtcct 485  
|||||

Db 417 TGCCCTTGTGACTGGCGAGTGGGGCCGAGCCACAAAGCGGAAGGTTCCGCTCCT 476  
|||||

QY 486 ggctgatccactgggcttggaaagagacagacttattactagatgattcggt 545  
|||||

Db 477 GGTGATCCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGTCGTGGT 536  
|||||

QY 546 gtccatctttgggaatgacgtctcaagaggttctccatggtgtacagatgacatag 605  
|||||

Db 537 GTCCATCTTTGGGAATGACGCTCAAGAGGTCTCCATGGTGTACAGATGTCATAGT 596  
|||||

QY 606 gaagccctgaatggaacagatggcagaagcctcacctgcagcctggcaccacaat 665  
|||||

Db 597 GAAGCCCTGATGTGAACAGATGGCACAGGCTCACCTGCAGCCTGGCACCAATAT 656  
|||||

QY 666 catctcacagctctgag 682  
|||||

Db 657 CATCTCACAGCTCTGAG 673  
|||||

RESULT 7  
AF197952 780 bp mRNA PRI 29-FEB-2000  
LOCUS Homo sapiens thioredoxin peroxidase PMP20 mRNA, complete cds.

AF197952  
AF197952.1 GI:6166492  
human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 780)  
Zhou, Y., Kok, K.H., Chun, A.C., Wong, C.M., Wu, H.W., Lin, M.C.,  
Fung, P.C., Kung, H. and Jin, D.Y.  
TITLE Mouse peroxiredoxin V is a thioredoxin peroxidase that inhibits  
p53-induced apoptosis  
JOURNAL Biochem. Biophys. Res. Commun. 268 (3), 921-927 (2000)  
MEDLINE 20145535  
REFERENCE 2 (bases 1 to 780)  
Jin, D.-Y.  
AUTHORS Direct Submission  
TITLE Submitted (23-OCT-1999) Institute of Molecular Biology, The  
JOURNAL University of Hong Kong, 8 Sassoon Road, Hong Kong  
FEATURES  
Location/Qualifiers  
1. .780  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11q13"  
110. .754  
/note="type V peroxiredoxin"  
/codon\_start=1  
/product="thioredoxin peroxidase PMP20"  
/protein\_id="AAF04856.1"  
/db\_xref="GI:6166493"  
/translation="MGLAGVLCALRRSAGVILVGGAGGQSAARRCSEGEWASGGVSR  
SFSRAAAMAPIKVGDAPFAVEFEGEPGNKVNLAELFKGKGVLFQVPGAFTPGCKS  
THLPFVEQEAALKAGVQVVAFLSYNDAPVTGEMGRAHKAEGKRVLLADPTGAFGKE  
TDLLDDSLVSIFGNRLKRFMSVMVDGIVKALNVEPDGTGLTCLAPNIISQL"

BASE COUNT 147 a 203 c 278 g 152 t

ORIGIN

Query Match 65.6%; Score 528; DB 89; Length 780;  
Best Local Similarity 99.8%; Pred. No. 3.5e-294;  
Matches 678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gccaggaggcggagtggaagtgccgtggcgcggtgatgggactagctgcgtgtgcgc 60  
|||||

Db 74 GCCAGGAGGCGGAGTGAAGTGGCGGTGGGGGGTATGGGACTAGCTGGCGTGGCGC 133  
|||||

QY 61 ctgagacgtcagcggggtatatactctgtgtggggcgcggtcagtcgtcggcgcg 120  
|||||

Db 134 CTGAGACGCTACGCGGGCTATATACTCGTGGTGGGGCGGGGTCAGTCTCGGCGAGCG 193  
|||||

QY 121 gcagcaagacggtgcagtgaaaggagagtggtggtctggcggttcgcagtttcagcaga 180  
|||||

Db 194 GCAGCAAGACGGTGCAGTGAAGGAGAGTGGGCTCTGGCGGGGTCGCGAGTTCAGCAGA 253  
|||||

QY 181 gccgtgcgccaatgcccccaatcaagtgggagatgccatcccgagcagtgaggtgttt 240  
|||||

Db 254 GCGCTGCAGCCATGCCCCCAATCAAGGTGGGAGATGCCATCCAGCAGTGGAGTGT 313  
|||||

QY 241 gaagggagccaggaacaagtggaacctggcagagctgttcaagggaagaggtgtg 300  
|||||

Db 314 GAAGGGACCCAGGNAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGAGGTGTG 373  
|||||

QY 301 ctgtttggagttcctggggccttcaccccttgatgttccaaagacacacctgcaggttt 360  
|||||

Db 374 CTGTTTGGAGTTCCTGGGCGCTTACCCTCGATGTTCCAAAGACACACCTGCCAGGTTT 433  
|||||

QY 361 gtggagcagctgagctctgaagccaaaggagtcacaggtgtgtgtgtgtgtgtgt 420  
|||||

Db 434 GTGGAGCAGGCTGAGGCTCTGAAGGCCCAAGGAGTCCAGGTGGTGGCTGTCTGAGTGT 493  
|||||

QY 421 aatgatgctttgtgactggcgagtggtggcgagagcccaagcggaaggaaggttcgg 480  
|||||

```

Db 494 AATGATGCTTTGTACTGGCGAGTGGGCGGAGCCACAAAGGCGGAAGGCAAGGTTCCG 553
QY 481 ctctggctgacccactgggccccttgggaagagacagacttattactagatgattcg 540
Db 554 CTCCTGGCTGATCCACATGGGCGCTTTTGGGAAGGAGACAGACTTATTACTAGATGATTCG 613
QY 541 ctgggtccatcttgggaatcagctctcaagaggttctccatgggtggttacagagatggc 600
Db 614 CTGGGTCTCATCTTTGGGAATCGACGCTCTCAAGAGGTTCTCCATGGGTGACAGATGGC 673
QY 601 atagtgaaggccctgaatgtgaaacagatggcacagcctcaactcagcctgagcctggcacc 660
Db 674 ATAGTGAAGGCCCTGAATGTGGAACAGATGSCACAGGCCCTCACCTGACGCTGGCACCC 733
QY 661 aatatactcacagctctga 681
Db 734 AATATCATCTCACAGCTCTGA 754

RESULT 8
HSA249483 489 bp mRNA PRI 02-DEC-1999
LOCUS Homo sapiens mRNA for human thiol peroxidase homologous protein
DEFINITION (TPX VI gene).
ACCESSION AJ249483
VERSION AJ249483.1 GI:6523288
KEYWORDS human thiol peroxidase homologous protein; TPX VI gene.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Kim, I. H. and Jeong, W.
JOURNAL A new type of human thiol peroxidase (Human TPx Type VI)
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 489)
TITLE Kim, I. H.
JOURNAL Direct Submission
SUBMITTED (13-SEP-1999) Kim I. H., Biochemistry, Paichai University,
439-6 Doma-2-Dong Seo-GU, Taejon 302-735, REPUBLIC OF KOREA
FEATURES
source
1. .489
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HeLa Cell"
1. 489
/gene="TPX VI"
1. .489
/gene="TPX VI"
/function="peroxidase"
/codon_start=1
/evidence="experimental"
/product="human thiol peroxidase homologous protein"
/protein_id="CAB62210.1"
/db_xref="GI:6523288"
/translation="MAPIKVGDAIPAVEFEGEPGNKNVLAELFKGKGVLFQVPGAF
TPGCSKTHLPFGVEQAEALAKAGQVQVACLVSNDAFVTGEWGRAHKAEGKVRLLADPT
GAFGKETDLLDDSLVSIFGNRLKRFMSVVQDGIKALNVEPDGTLTCSLAPNIIS
QL"
BASE COUNT 107 a 114 c 163 g 105 t
ORIGIN

Query Match 60.7%; Score 489; DB 93; Length 489;
Best Local Similarity 100.0%; Pred. No. 1.4e-271;
Matches 489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 atggccccaatcaaggtggagatgccatccagcagtgagggtgtttgaagggagacca 252
Db 1 ATGGCCCCAATCAAGGTGGGAGATGCCATCCAGCAGTGAGGTGTGTGAAGGGGAGCCA 60
QY 253 gggacaagggtgaacctggcagactgttcaaggcgaaggggtgtgctgtttgagtt 312

```

```

Db 61 GGAACAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGAGGGTGTGCTGTTGGAGTT 120
QY 313 cctgggccccttacccttgatgttccaagacacacctgccagggtttgtgagcaggct 372
Db 121 CTTGGGGCTTACCCCTGGATGTTCCAGACACACACTGCCAGGGTTTGTGGAGCAGGCT 180
QY 373 gaggtctgaagcccaagggagtcctcagggtgggtggtcctgtctgagtgatgaatgagccttt 432
Db 181 GAGGCTCTGAAGGCCAAGGAGTCCAGGTGGTGGCCCTGTCTGAGTGTATAATGATGCTTT 240
QY 433 gtactggcagatggggccagcccaagcggaaggaaggttcggctcctcggctgat 492
Db 241 GTGACTGGCGAGTGGGGCCGAGCCACAGCGGAGGACAGACTTATTACTAGATGATTCGCTGGCTGAT 300
QY 493 cccactgggccccttgggaagggagacagacttattactagatgattcgtcgtggtccatc 552
Db 301 CCCACTGGGCCITTTGGGAAGGAGACAGACTTATTACTAGATGATTCGCTGGTGTCCATC 360
QY 553 ttgggaatcgacgtctcgaagaggttccatggttggttacaggatggcatagtaagggc 612
Db 361 TTTGGGAATCGACGCTCTCAAGAGGTTCTCCATGGTGTGATGAGGATGATGATGAGGCC 420
QY 613 ctgaatgaaacagatggcacagcctcacctgcagcctggcaccacataatcatctca 672
Db 421 CTGAATGTGGAACCATGATGGCACAGGCTCACCTGACGCTGGCACCCCAATATCATCTCA 480
QY 673 cagctctga 681
Db 481 CAGCTCTGA 489

RESULT 9
A99104 601 bp DNA PAT 20-SEP-2000
LOCUS Sequence 8 from Patent WO9909054.
DEFINITION A99104
ACCESSION A99104
VERSION A99104.1 GI:6782055
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 601)
TITLE Falmagne, P., Wattiez, R., Bernard, A., Hermans, C. and Knoops, B.
PEROXISOME-ASSOCIATED POLYPEPTIDE, NUCLEOTIDE SEQUENCE ENCODING
SAID POLYPEPTIDE AND THEIR USES IN THE DIAGNOSIS AND/OR THE
TREATMENT OF LUNG INJURIES AND DISEASES, AND OF OXIDATIVE
STRESS-RELATED DISORDERS
JOURNAL Patent: WO 9909054-A 8 25-FEB-1999;
UNIV MONS HAINAUT (BE); FALMAGNE PAUL (BE); WATTIEZ RUDDY (BE);
BERNARD ALFRED (BE); HERMANS CEDRIC (BE); KNOOPS BERNARD (BE);
UNIV LOUVAIN (BE)
FEATURES
source
1. .601
/organism="Homo sapiens"
/db_xref="taxon:9606"
161. .514
/notice="unnamed protein product"
/codon_start=1
/protein_id="CAB70593.1"
/db_xref="GI:6782056"
/translation="MAPIKTHLPFGVEQAEALAKAGQVQVACLVSNDAFVTGEWGRAH
KAEGKVRLLADPTGAFGKETDLLDDSLVSIFGNRLKRFMSVVQDGIKALNVEPDG
TGLTCSLAPNIISQL"
BASE COUNT 115 a 162 c 200 g 124 t
ORIGIN

Query Match 53.4%; Score 430; DB 9; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.1e-237;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 339 caagcacacacctgccagggtttgtgagcaggtgaggtcgtgaaggccaagggagttcca 398

```

[illegible]

```

RESULT 10
A99105
LOCUS      604 bp      DNA
DEFINITION      Sequence 9 from Patent WO9909054.
ACCESSION      A99105
VERSION
KEYWORDS
SOURCE
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 604)
Falmagne,P., Wattiez,R., Bernard,A., Hermans,C. and Knoops,B.
Peroxisome-associated polypeptide, nucleotide sequence encoding
said polypeptide and their uses in the diagnosis and/or the
treatment of lung injuries and diseases, and of oxidative
stress-related disorders
Patent: WO 9909054-A 9 25-FEB-1999;
UNIV MONS HAINAUT (BE) ; FALMAGNE PAUL (BE) ; WATTIEZ RUDDY (BE) ;
BERNARD ALFRED (BE) ; HERMANS CEDRIC (BE) ; KNOOPS BERNARD (BE) ;
UNIV LOUVAIN (BE)
FEATURES
    source          Location/Qualifiers
        1..604
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
        161..517
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAB70594.1"
            /db_xref="GI:6782058"
            /translation="MAPIKVGDPAIVEFEGEPGNKVNLAELFKGKGVLPFCVPGAF
            TGGLTCSLAPNIISOL"
            /catalytic_activity="GTPase"
BASE COUNT      117 a      160 c      200 g      127 t
ORIGIN

```

```

Query Match          38.5%; Score 310; DB 9; Length 604;
Best Local Similarity 100.0%; Pred. No. 7e-168;
Matches 310; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY	33	gggtatgggactagctggcggtgctgcgcctgagacgctcagcgggctatatactcgtcg	92
Db	1	GGGTATGGGACTAGCTGGCGGTGTCGCGCTGAGACGCTCAGCGGGCTATATACTTCGTGG	60
QY	93	tggggccggcgctcagctcgcggcagcgagcaagaagctgcaagtgaagagtggtggc	152
Db	61	TGGGGCCGCGGTCAGTCTCGCGCAGCGGAGCAAGACGCTGCAGTGAAGGAGAGAGTGGC	120
QY	153	gtctggcggggtccgcagtttcagcagagcgcgtgcagccatggccccaatcaaggtggg	212
Db	121	GTCTGGCGGGGTCCGAGTTTCAGCAGAGCGCGTGCAGCCATGCCCCATCAAGTGGG	180
QY	213	agatgccatccacagtcagtgaggtgtttaaagggagccagggaaacaagtggaacctggc	272
Db	181	AGATGCCATCCACAGCAGTGGAGGTGTTGAAGGGGAGCCAGGGAAAGGTGAACCTGGC	240
QY	273	agagctgtccaagggcaagaaggtgtgctgtttgagttctctggggcgcttcaccctctg	332
Db	241	AGAGCTGTTCAGGGCAAGAGGGTGTGCTGTTTGGAGTTCCTGGGGCCTTCACCCCTGG	300
QY	333	atgttccaag 342	
Db	301	ATGTTCGAAG 310	

RESULT	11
A99103	
LOCUS	A99103 469 bp DNA
DEFINITION	Sequence 7 from Patent WO9909054.
ACCESSION	A99103
VERSION	A99103.1 GI:6782053
KEYWORDS	

RESULT 11

A99103		DNA	PAT	20-SEP-2000
LOCUS	A99103	469 bp		
DEFINITION	Sequence 7 from Patent WO9909054.			
ACCESSION	A99103			
VERSION	A99103.1	GI:6782053		
KEYWORDS	.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 469)			
AUTHORS	Falmane,P., Wattiez,R., Bernard,A., Hermans,C. and Knoops,B.			
TITLE	Peroxisome-associated polypeptide, nucleotide sequence encoding said polypeptide and their uses in the diagnosis and/or the treatment of lung injuries and diseases, and of oxidative stress-related disorders			
JOURNAL	PATENT : WO 9909054-A 7 25-FEB-1999;			
	UNIV MONS HAINAUT (BE) ; FALMAGNE PAUL (BE) BERNARD ALFRED (BE) ; HERMANNS CEDRIC (BE) ; KNOOPS BERNARD (BE) ; UNIV LOUVAIN (BE)			

## FEATURES source

```

1: .403
/organism="Homo sapiens"
/db_xref="taxon:9606"
161.382
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB70592.1"
/db_xref="GI:6782054"
/translation="MAPIKVRLIADPTGAFKETDLLDSLVIFGNRRRLKRFMSVV
QDGIVKALNVEPDGTLGTCSLAPNISQL"
88 a 132 c 150 q 99 t
BASE COUNT

```

```
Query Match          32.7%; Score 263; DB 9; Length 469;
Best Local Similarity 100.0%; Pred. No. 1.2e-140;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	471	caaggttcggtcctcggtcgatcccatcgggccttgggaaggagacagacttatct	530
Db	172	CAAGGTTGGCTCCTGCTCATCCACTGGGGCCTTTGGGAAGGAGACAGACTTATTA	231
Qy	531	agatgattcggtgtcccatctttgggaatcgagctctcaagaggtctcccatgggggt	590
Db	232	AGATGATTTCGCTGGTGCCATCTTTGGGAATCGACGCTCTCAAGAGGTTCTCCATGGTGGT	291

\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 701: contig of 701 bp in length  
702 801: gap of 100 bp  
802 1496: contig of 695 bp in length  
1497 1596: gap of 100 bp  
1597 2252: contig of 656 bp in length  
2253 2352: gap of 100 bp  
2353 3061: contig of 709 bp in length  
3062 3161: gap of 100 bp  
3162 3869: contig of 708 bp in length  
3870 3969: gap of 100 bp  
3970 4681: contig of 712 bp in length  
4682 4781: gap of 100 bp  
4782 5514: contig of 733 bp in length  
5515 5614: gap of 100 bp  
5615 6299: contig of 685 bp in length  
6300 6399: gap of 100 bp  
6400 7087: contig of 688 bp in length  
7088 7187: gap of 100 bp  
7188 7862: contig of 675 bp in length  
7863 7962: gap of 100 bp  
7963 8673: contig of 711 bp in length  
8674 8773: gap of 100 bp  
8774 9484: contig of 711 bp in length  
9485 9584: gap of 100 bp  
9585 10302: contig of 718 bp in length  
10303 10402: gap of 100 bp  
10403 11104: contig of 702 bp in length  
11105 11204: gap of 100 bp  
11205 11921: contig of 717 bp in length  
11922 12021: gap of 100 bp  
12022 12698: contig of 677 bp in length  
12699 12798: gap of 100 bp  
12799 13513: contig of 715 bp in length  
13514 14327: contig of 714 bp in length  
14328 14427: gap of 100 bp  
14428 15106: contig of 679 bp in length  
15107 15206: gap of 100 bp  
15207 15877: contig of 671 bp in length  
15878 15977: gap of 100 bp  
15978 16670: contig of 693 bp in length  
16671 16770: gap of 100 bp  
16711 17443: contig of 673 bp in length  
17444 17543: gap of 100 bp  
17544 18233: contig of 690 bp in length  
18234 18333: gap of 100 bp  
18334 19012: contig of 679 bp in length  
19013 19112: gap of 100 bp  
19113 19806: contig of 694 bp in length  
19807 19906: gap of 100 bp  
19907 20573: contig of 667 bp in length  
20574 20673: gap of 100 bp  
20674 21371: contig of 698 bp in length  
21372 21471: gap of 100 bp  
21472 22184: contig of 713 bp in length  
22185 22284: gap of 100 bp  
22285 22976: contig of 692 bp in length  
22977 23076: gap of 100 bp  
23077 23749: contig of 673 bp in length  
23750 23849: gap of 100 bp  
23850 24526: contig of 677 bp in length  
24527 24626: gap of 100 bp  
24627 25319: contig of 693 bp in length  
25320 25419: gap of 100 bp  
25420 26077: contig of 658 bp in length  
26078 26177: gap of 100 bp  
26178 26838: contig of 661 bp in length  
26839 26938: gap of 100 bp  
26939 27644: contig of 706 bp in length

QY 591 acagatgcatagtgaggccctgaatgtggaaccagatggcacaggcctcacctgcag 650  
|||||  
Db 292 ACAGGATGCGATAGTGAAGGCCCTGAATGTGGAACACAGATGGCACAGCCCTCACCTGCAG 351  
|||||

QY 651 cctggcaccacaatactatctacagctctgagcgccctggccagatgattctctctccacccc 710  
|||||  
Db 352 CTTGGCACCACCATATCATCTACAGCTCTGAGGCCCTGGCCAGATTACTTCTCTCCACCC 411  
|||||

QY 711 ctccctatctacccctgcagccc 733  
|||||  
Db 412 CTCCCTATCTACCTGCCAGCC 434  
|||||

RESULT 12  
AC090391 49259 bp DNA HTG 17-FEB-2001  
LOCUS Homo sapiens chromosome 11 clone RP11-87716 map 11, LOW-PASS  
DEFINITION SEQUENCE SAMPLING.

AC090391  
VERSION AC090391.1 GI:12958027  
KEYWORDS HTG; HTGS\_PHASE0.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
1 (bases 1 to 49259)  
TITLE Homo sapiens chromosome 11, clone RP11-87716  
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 49259)  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,  
Barn,N., Bastien,V., Boguslavskiy,L., Boukhalter,B., Brown,A.,  
Cammarata,J., Campopiano,A., Choepel,V., Colangelo,M., Collins,S.,  
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,  
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Hearford,A., Horton,L., Hulme,W., Illiev,I., Johnson,R.,  
Jones,C., Karatas,A., Larocque,K., Lamazares,R., Landers,T.,  
Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,  
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,  
McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,  
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,  
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,  
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Travers,M., Travis,N., Trigglio,J., Vassiliev,H., Viel,R., Vo,A.,  
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission  
JOURNAL Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
COMMENT All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L12693  
Center clone name: 877\_I\_6  
-----

\* NOTE: This record contains 62 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.



Mon Oct 22 14:46:45 2001

27739 31508 contig of 3770 bp in length  
31609 35096 contig of 3488 bp in length  
35197 39266 contig of 4070 bp in length  
39367 41973 contig of 2607 bp in length  
42074 43566 contig of 1493 bp in length  
43667 46358 contig of 2692 bp in length  
46459 49035 contig of 2577 bp in length  
49136 50654 contig of 1519 bp in length  
50755 53342 contig of 2588 bp in length  
53443 56951 contig of 1792 bp in length  
55335 56951 contig of 1617 bp in length  
57052 59082 contig of 2031 bp in length  
59183 60880 contig of 1698 bp in length  
60981 62607 contig of 1627 bp in length  
62708 63786 contig of 1079 bp in length  
63887 64901 contig of 1015 bp in length  
65002 66266 contig of 1265 bp in length

Sequence updated (04-Nov-1999)  
Sequence updated (26-May-2000).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 23 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
6017: contig of 6017 bp in length  
6018 6117: gap of 100 bp  
6118 11749: contig of 5632 bp in length  
11750 11849: gap of 100 bp  
11850 16292: contig of 4443 bp in length  
16293 16392: gap of 100 bp  
16393 19696: contig of 3304 bp in length  
19697 19796: gap of 100 bp  
19797 23376: contig of 3580 bp in length  
23377 23476: gap of 100 bp  
23477 27638: contig of 4162 bp in length  
27639 27738: gap of 100 bp  
27739 31508: contig of 3770 bp in length  
31509 31608: gap of 100 bp  
31609 35096: contig of 3488 bp in length  
35097 35196: gap of 100 bp  
35197 39266: contig of 4070 bp in length  
39267 39366: gap of 100 bp  
39367 41973: contig of 2607 bp in length  
41974 42073: gap of 100 bp  
42074 43566: contig of 1493 bp in length  
43567 43666: gap of 100 bp  
43667 46358: contig of 2692 bp in length  
46359 46458: gap of 100 bp  
46459 49035: contig of 2577 bp in length  
49036 49135: gap of 100 bp  
49136 50654: contig of 1519 bp in length  
50655 50754: gap of 100 bp  
50755 53342: contig of 2588 bp in length  
53343 53442: gap of 100 bp  
53443 55234: contig of 1792 bp in length  
55235 55334: gap of 100 bp  
55335 56951: contig of 1617 bp in length  
56952 57051: gap of 100 bp  
57052 59082: contig of 2031 bp in length  
59083 59182: gap of 100 bp  
59183 60880: contig of 1698 bp in length  
60881 60980: gap of 100 bp  
60981 62607: contig of 1627 bp in length  
62608 62707: gap of 100 bp  
62708 63786: contig of 1079 bp in length  
63787 63886: gap of 100 bp  
63887 64901: contig of 1015 bp in length  
64902 65001: gap of 100 bp  
65002 66266: contig of 1265 bp in length.

Location/Qualifiers

source

1. .56266  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11q13"  
/clone="XXp1-25H10"  
1. .6017  
/note="assembly\_fragment"  
6118. .11749  
/note="assembly\_fragment"  
11850. .16292  
/note="assembly\_fragment"  
16393. .19696  
/note="assembly\_fragment"  
19797. .23376  
/note="assembly\_fragment"  
23477. .27638  
/note="assembly\_fragment"  
27739. .31508  
/note="assembly\_fragment"  
31609. .35096  
/note="assembly\_fragment"  
35197. .39266  
/note="assembly\_fragment"  
39367. .41973  
/note="assembly\_fragment"  
42074. .43566  
/note="assembly\_fragment"  
43667. .46358  
/note="assembly\_fragment"  
46459. .49035  
/note="assembly\_fragment"  
49136. .50654  
/note="assembly\_fragment"  
50755. .53342  
/note="assembly\_fragment"  
53443. .55234  
/note="assembly\_fragment"  
55335. .56951  
/note="assembly\_fragment"  
57052. .59082  
/note="assembly\_fragment"  
59183. .60880  
/note="assembly\_fragment"  
60981. .62607  
/note="assembly\_fragment"  
62708. .63786  
/note="assembly\_fragment"  
63887. .64901  
/note="assembly\_fragment"  
65002. .66266  
/note="assembly\_fragment"  
BASE COUNT 13932 a 17688 c 18127 g 14319 t 2200 others  
ORIGIN

Query Match 28.0%; Score 225; DB 82; Length 66266;  
Best Local Similarity 100.0%; Pred. No. 1.2e-118; Indels 0; Gaps 0;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 aggttcctcatggtacagatgtagtgaagccctgaatgtggaaccagatgac 633  
|||||  
Db 13048 AGGTTCTCCATGTTGTACAGATGCGATGATGAGGCCCTGATGTGACACAGATGCG 13107  
QY 634 acaggcctcacctgcagcctggcaccacaataatcatcatcagctctgagccctggcca 693  
|||||  
Db 13108 ACAGGCTTCACCTGCAGCTGGCACCAATATCATCTCAGCTCTGAGGCCCTGGGCA 13167  
QY 694 gattacttctcacccctccctcccttatctcaccctggccagccctgtgctgggcccctgcaat 753  
|||||  
Db 13168 GATTACTTCTCCACCCCTCCCTTATCTACCTGCGCCGCTGTGCTGGGCCCTGCAAT 13227  
QY 754 tggaaatgtggccagattttctgcaataaacactgtggttggcgg 798

Db 13228 TGGATGTGCCAGATTCTGCAATAACACTGTGTTGCGG 13272  
 |||

## RESULT 14

AP001453

LOCUS

DEFINITION

AP001453 151747 bp DNA HTG 14-SEP-2000  
 Homo sapiens chromosome 11 clone RP11-783K16 map 11q13, WORKING  
 DRAFT SEQUENCE, 37 unordered pieces.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

## COMMENT

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 37 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1  
 \* 8903 9002: contig of 8902 bp in length  
 \* 9003 18232: gap of 100 bp  
 \* 18233 18332: contig of 9230 bp in length  
 \* 18333 28659: contig of 10327 bp in length  
 \* 28660 28759: gap of 100 bp  
 \* 28760 37331: contig of 8572 bp in length  
 \* 37332 37431: gap of 100 bp  
 \* 37432 45066: contig of 7635 bp in length  
 \* 45067 45166: gap of 100 bp  
 \* 45167 51737: contig of 6571 bp in length  
 \* 51738 51837: gap of 100 bp  
 \* 51838 57917: contig of 6080 bp in length  
 \* 57918 58017: gap of 100 bp  
 \* 58018 62697: contig of 4680 bp in length  
 \* 62698 62797: gap of 100 bp  
 \* 62798 66342: contig of 3545 bp in length  
 \* 66343 66442: gap of 100 bp  
 \* 66443 69246: contig of 2804 bp in length  
 \* 69247 69346: gap of 100 bp  
 \* 69347 72173: contig of 2827 bp in length  
 \* 72174 72273: gap of 100 bp  
 \* 72274 76432: contig of 4159 bp in length  
 \* 76433 76532: gap of 100 bp  
 \* 76533 80848: contig of 4316 bp in length  
 \* 80849 80948: gap of 100 bp  
 \* 80949 84529: contig of 3581 bp in length  
 \* 84530 84629: gap of 100 bp  
 \* 84630 88588: contig of 3959 bp in length  
 \* 88589 88689: gap of 100 bp  
 \* 88689 94224: contig of 5536 bp in length  
 \* 94225 94324: gap of 100 bp  
 \* 94325 98727: contig of 4403 bp in length  
 \* 98728 98827: gap of 100 bp  
 \* 98828 102161: contig of 3333 bp in length  
 \* 102161 102260: gap of 100 bp  
 \* 102261 107715: contig of 5455 bp in length  
 \* 107716 107815: gap of 100 bp  
 \* 107816 111480: contig of 3665 bp in length  
 \* 111481 111580: gap of 100 bp  
 \* 111581 114035: contig of 2455 bp in length  
 \* 114036 114135: gap of 100 bp  
 \* 114136 117720: contig of 3585 bp in length  
 \* 117721 117821: gap of 100 bp  
 \* 117821 121176: contig of 3356 bp in length

Center: RIKEN Genomic Sciences Center (GSC)  
 Center code: RIKEN  
 Web site: <http://hgp.gsc.riken.go.jp/>  
 Contact: [hattori@gsr.riken.go.jp](mailto:hattori@gsr.riken.go.jp)  
 ----- Project Information  
 Center project name: Humdraft11  
 Center clone name: RP11-783K16  
 ----- Summary Statistics  
 Sequencing vector: PCR products; 100% of reads  
 Chemistry: Dye-terminator ET-amersham; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 136959 bases at least Q40  
 Consensus quality: 143151 bases at least Q30  
 Consensus quality: 146210 bases at least Q20  
 Insert size: 148147; sum-of-contigs  
 Quality coverage: 8.32x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of  
 37 contigs. The true order of the pieces is not known and their  
 order in this sequence record is arbitrary. Gaps between the  
 contigs are represented as runs N, but the exact sizes of the gaps  
 are unknown. This record will be updated with the finished sequence  
 as soon as it is available and the accession number will be  
 preserved

1  
 \* 8902 contig of 8902 bp in length  
 \* 9003 18232: contig of 9230 bp in length  
 \* 18333 28659: contig of 10327 bp in length  
 \* 28760 37331: contig of 8572 bp in length  
 \* 37432 45066: contig of 7635 bp in length  
 \* 45167 51737: contig of 6571 bp in length  
 \* 51838 57917: contig of 6080 bp in length  
 \* 58018 62697: contig of 4680 bp in length  
 \* 62798 66342: contig of 3545 bp in length  
 \* 66443 69246: contig of 2804 bp in length  
 \* 69347 72173: contig of 2827 bp in length  
 \* 72274 76432: contig of 4159 bp in length  
 \* 76533 80848: contig of 4316 bp in length  
 \* 80949 84529: contig of 3581 bp in length  
 \* 84630 88588: contig of 3959 bp in length  
 \* 88689 94224: contig of 5536 bp in length  
 \* 94325 98727: contig of 4403 bp in length

Qy 694 gattactctccacccctccctctatctccactgcccagccctgtgtggccctgcaat 753  
Dy 48315 GATTACTTCTCCACCCCTCCCTATCTCAGCTGCCAGCCCTGTGTGGGGCCCTGCAAT 48374  
Qy 754 tggaatgtggccagatttctgcaataaacacattgtggttgcgg 798  
Dy 48375 TGGAAATGTGGCCAGATTCTGCAATAAACACATTGTGGTTTTCGGG 48419

RESULT 15  
AC015698/168475 bp DNA HTG 22-OCT-2000  
AC015698/168475 bp DNA HTG 22-OCT-2000  
LOCUS Homo sapiens chromosome 11 clone RP11-147G6 map 11, WORKING DRAFT  
DEFINITION SEQUENCE, 9 unordered pieces.  
ACCESSION AC015698  
VERSION AC015698.4 GI:10945776  
KEYWORDS HIG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 168475)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,  
Brown, A., Castle, A., Collinge, R., Collins, S., Collymore, A.,  
Cooke, P., Dearlano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,  
Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D.,  
Galagan, J., Gardyna, S., Grant, G., Hagg, B., Heaford, A., Horton, L.,  
Howland, J., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J.,  
Lehoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,  
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,  
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Oct 22, 2000 this sequence version replaced gi:7341708.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L1368  
Center clone name: L1368  
----- Summary Statistics  
Sequencing vector: M13; M77815; 41% of reads  
Sequencing vector: plasmid; n/a; 59% of reads  
Chemistry: Dye-primer-amersham; 0% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 165654 bases at least Q40  
Consensus quality: 167047 bases at least Q30  
Consensus quality: 167433 bases at least Q20  
Insert size: 170000; agarose-fp  
Insert size: 167675; sum-of-contigs  
Quality coverage: 10.1 in Q20 base.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 9 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence

121177 121276: gap of 100 bp  
121277 123962: contig of 2686 bp in length  
123963 124062: gap of 100 bp  
124063 128424: contig of 4362 bp in length  
128425 128524: gap of 100 bp  
128525 131280: contig of 2756 bp in length  
131281 131380: gap of 100 bp  
131381 133764: contig of 2384 bp in length  
133765 133864: gap of 100 bp  
133865 137166: contig of 3302 bp in length  
137167 137266: gap of 100 bp  
137267 139568: contig of 2702 bp in length  
139569 140068: gap of 100 bp  
140069 141380: contig of 1312 bp in length  
141381 141480: gap of 100 bp  
141481 142715: contig of 1235 bp in length  
142716 142815: gap of 100 bp  
142816 144247: contig of 1432 bp in length  
144248 144347: gap of 100 bp  
144348 146022: contig of 1675 bp in length  
146023 146122: gap of 100 bp  
146123 147416: contig of 1294 bp in length  
147417 147516: gap of 100 bp  
147517 148926: contig of 1410 bp in length  
148927 149026: gap of 100 bp  
149027 150172: contig of 1146 bp in length  
150173 150272: gap of 100 bp  
150273 151747: contig of 1475 bp in length.

FEATURES  
Location/Qualifiers  
1..151747  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11q13"  
/clone="RP11-783K16"  
1..8902  
/note="assembly\_fragment"  
9003..18232  
/note="assembly\_fragment"  
18333..28659  
/note="assembly\_fragment"  
28760..37331  
/note="assembly\_fragment"  
37432..45066  
/note="assembly\_fragment"  
45167..51737  
/note="assembly\_fragment"  
51838..57917  
/note="assembly\_fragment"  
58018..62697  
/note="assembly\_fragment"  
62798..66342  
/note="assembly\_fragment"  
66443..69246  
/note="assembly\_fragment clone\_end:r7 vector\_side:right"  
69347..72173  
/note="assembly\_fragment clone\_end:sp6 vector\_side:right"  
72274..76432  
/note="assembly\_fragment"  
76533..80848

Query Match 28.0%; Score 225; DB 82; Length 151747;  
Best Local Similarity 100.0%; Pred. No. 1.2e-118;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 574 aggtctccatggtgtgacagatggtgagccctgaatggaaccagatggc 633  
Dy 48195 AGGTTCTCCATGGTGTACAGATGGGCATAGTGAAGCCCTGATGTTGGACACAGATGGC 48254  
Qy 634 acagcctccactgagcctggcccaataatcatctcacagctctgagccctggcca 693  
Dy 48255 ACAGGCTCACCCTGAGCCTGGCACCACCAATATCATCTCACAGCTCTGAGGCCCTGGGCCA 48314

\* as soon as it is available and the accession number will  
 \* be preserved.

1 40714: contig of 40714 bp in length  
 40715 40814: gap of 100 bp  
 40815 42233: contig of 1419 bp in length  
 42234 42333: gap of 100 bp  
 42334 47280: contig of 4947 bp in length  
 47281 47380: gap of 100 bp  
 47381 53604: contig of 6224 bp in length  
 53605 53704: gap of 100 bp  
 53705 61923: contig of 8225 bp in length  
 61930 62023: gap of 100 bp  
 62030 109187: contig of 47158 bp in length  
 109188 109287: gap of 100 bp  
 109288 131160: contig of 21873 bp in length  
 131161 131260: gap of 100 bp  
 131261 150366: contig of 19106 bp in length  
 150367 150466: gap of 100 bp  
 150467 168475: contig of 18009 bp in length.

## FEATURES

source

1. 168475  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="11"  
 /map="11"  
 /clone="rp11-147G6"  
 /clone\_lib="RPCI-11 Human Male BAC"  
 1. 40714  
 /note="assembly\_fragment"  
 clone\_end:SP6  
 vector\_side:left  
 40815..42233  
 /note="assembly\_fragment"  
 42334..47280  
 /note="assembly\_fragment"  
 47381..53604  
 /note="assembly\_fragment"  
 53705..61929  
 /note="assembly\_fragment"  
 62030..109187  
 /note="assembly\_fragment"  
 109288..131160  
 /note="assembly\_fragment"  
 131261..150366  
 /note="assembly\_fragment"  
 150467..168475  
 /note="assembly\_fragment"  
 clone\_end:17  
 vector\_side:right"

BASE COUNT 40113 a 42972 c 43226 g 41363 t 801 others  
 ORIGIN

Query Match 28.0%; Score 225; DB 63; Length 168475;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-118;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 574 aggttcctccatggttgacagatggcatagtgagccctgaatgggaaccagatggc 633  
 |||||  
 Db 165891 AGGTTCTCCATGGTGTACAGGATGGCATAGTGAAGGCCCTGGAATGTGAACACAGATGGC 165832  
 QY 634 acagccctcactgcagcctggcaccacatcatctcacagctctgagccctgggcca 693  
 |||||  
 Db 165831 ACAGCCCTCACCTGCAGCCTGGCACCACCAATATCATCTCACAGCTCTGAGGCCCTGGCCA 165772  
 QY 694 gattactctcccccctccctctatctcactgcagccctgctgctggggccctgcaat 753  
 |||||  
 Db 165771 GATTACTTCTCTCCACCCCTCCCTATCTACCTGCCACAGCCCTGTGTGGGCCCTGCAAT 165712  
 QY 754 tggatgttgccagatttctgcaataaacacactgtgtgttgcgg 798  
 |||||  
 Db 165711 TGGATGTGTGGCCAGATTCTTGTGAATAAACACATTGTGTTTGGGG 165667

Search completed: October 22, 2001, 13:44:51  
 Job time: 5461 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2001, 12:12:15 ; Search time 1291.82 seconds  
(without alignments)  
5890.563 Million cell updates/sec

Title: US-09-486-167A-1

Perfect score: 805

Sequence: 1 gccaggagcgagtggaag.....ttgtgttgcgaaaaaaa 805

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 10228115 seqs, 4726426750 residues

Word size: 0

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est33:\*  
26: gb\_est34:\*  
27: gb\_est35:\*  
28: gb\_est36:\*  
29: gb\_est37:\*  
30: gb\_est38:\*  
31: gb\_est39:\*  
32: gb\_est40:\*  
33: em\_estba:\*  
34: em\_estfun:\*  
35: em\_esthum1:\*  
36: em\_esthum2:\*  
37: em\_esthum3:\*  
38: em\_esthum4:\*  
39: em\_esthum5:\*  
40: em\_esthum6:\*  
41: em\_esthum7:\*  
42: em\_esthum8:\*  
43: em\_esthum9:\*  
44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_estin1:\*  
64: em\_estin2:\*  
65: em\_estin3:\*  
66: em\_estin4:\*  
67: em\_estin5:\*  
68: em\_estom1:\*  
69: em\_estom2:\*  
70: em\_estov1:\*  
71: em\_estov2:\*  
72: em\_estpl1:\*  
73: em\_estpl2:\*  
74: em\_estpl3:\*  
75: em\_estpl4:\*  
76: em\_estpl5:\*  
77: em\_estpl6:\*  
78: em\_estpl7:\*  
79: em\_estpl8:\*  
80: em\_estpl9:\*  
81: em\_estpl10:\*  
82: em\_estro1:\*  
83: em\_estro2:\*  
84: em\_estro3:\*  
85: em\_estro4:\*  
86: em\_estro5:\*  
87: em\_estro6:\*  
88: em\_estro7:\*  
89: em\_estro8:\*  
90: em\_estro9:\*  
91: em\_estro10:\*  
92: em\_estro11:\*  
93: em\_estro12:\*  
94: em\_estro13:\*  
95: em\_estro14:\*  
96: em\_estro15:\*  
97: em\_estro16:\*  
98: em\_estro17:\*  
99: em\_estro18:\*  
100: em\_estro19:\*  
101: em\_estro20:\*  
102: gb\_est25:\*  
103: gb\_est26:\*  
104: gb\_est27:\*  
105: gb\_est28:\*  
106: gb\_est29:\*  
107: gb\_est30:\*  
108: gb\_est31:\*  
109: gb\_est32:\*  
110: gb\_est41:\*  
111: gb\_est42:\*  
112: gb\_est43:\*  
113: gb\_est44:\*  
114: gb\_est45:\*  
115: gb\_est46:\*  
116: gb\_est47:\*



Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
Source

Location/Qualifiers  
1..918  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DE007YL04"  
/clone\_lib="LTI\_FL002\_PL1"  
/lab\_host="DH10B"  
/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand  
cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-stranded cDNA was digested with Not I  
and cloned into the Not I and Eco RV sites of the  
pCMVSPORT 6 vector. Library was constructed by Life  
Technologies. Contact : Feng Liang Life Technologies, a  
division of Invitrogen 9800 Medical Center Drive Rockville  
Maryland 20850, USA Fax : (1) 301 610 8371 Email :  
filiang@lifetech.com URL :  
http://fulllength.invitrogen.com"

BASE COUNT 218 a 225 c 296 g 178 t 1 others  
ORIGIN

Query Match 89.2%; Score 718; DB 106; Length 918;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 768; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gccagagcgagtggaagtggcggtggggcggtatggagctagctggcgtgccc 60  
DB 62 gccagagcgagtggaagtggcggtggggcggtatggagctagctggcgtgccc 121  
QY 61 ctgagacgctcagcggtctatactcgtcggtggcggtggcggtcagctcgtgcagcg 120  
DB 122 ctgagacgctcagcggtctatactcgtcggtggcggtggcggtcagctcgtgcagcg 181  
QY 121 gcagcaagcggtcagtggaagagagtggtggcggtcgtggcggtcagtggaagag 180  
DB 182 gcagcaagcggtcagtggaagagagtggtggcggtcgtggcggtcagtggaagag 241  
QY 181 gccgtcagcggtcagtggaagagagtggtggcggtcgtggcggtcagtggaagag 240  
DB 242 gccgtcagcggtcagtggaagagagtggtggcggtcgtggcggtcagtggaagag 301  
QY 241 gaaggagcggtcagtggaagagagtggtggcggtcgtggcggtcagtggaagag 300  
DB 302 gaaggagcggtcagtggaagagagtggtggcggtcgtggcggtcagtggaagag 361  
QY 301 ctgtttggagttcttggcggtcagtggaagagagtggtggcggtcagtggaagag 360  
DB 362 ctgtttggagttcttggcggtcagtggaagagagtggtggcggtcagtggaagag 421  
QY 361 gtggagcaggtcagtggaagagagtggtggcggtcagtggaagagagtggtggtg 420  
DB 422 gtggagcaggtcagtggaagagagtggtggcggtcagtggaagagagtggtggtg 481  
QY 421 aatgatgctttgtgactggcggtcagtggaagagagtggtggcggtcagtggaagag 480  
DB 482 aatgatgctttgtgactggcggtcagtggaagagagtggtggcggtcagtggaagag 541  
QY 481 ctctgtgctgactggcggtcagtggaagagagtggtggcggtcagtggaagagagtg 540  
DB 542 ctctgtgctgactggcggtcagtggaagagagtggtggcggtcagtggaagagagtg 601  
QY 541 ctggtgtcctatcttggggaatcagtggtcagtggaagagagtggtggcggtcagtg 600  
DB 602 ctggtgtcctatcttggggaatcagtggtcagtggaagagagtggtggcggtcagtg 661  
QY 601 atagtgaagcggtcagtggaagagagtggtggcggtcagtggaagagagtggtggcg 660  
DB 662 atagtgaagcggtcagtggaagagagtggtggcggtcagtggaagagagtggtggcg 721  
QY 661 aatatcatctcacagctcagtggtggcggtcagtggaagagagtggtggcggtcagtg 720

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	718	89.2	918	106	AL541916	AL541916
2	641	79.6	677	32	AV650400	AV650400
3	633	78.6	790	140	BE784530	BE784530
4	622	77.3	622	175	BE255777	BE255777
5	622	77.3	703	175	BE287024	BE287024
6	612	76.0	840	172	BE026788	BE026788
7	604	75.0	604	165	BE298055	BE298055
8	585	72.7	1064	172	BE976165	BE976165
9	584	72.5	959	167	BE410064	BE410064
10	583	72.4	831	154	BE498369	BE498369
11	572	71.1	623	172	BE971772	BE971772
12	568	70.6	818	106	AL535869	AL535869
13	550	68.3	710	152	BE325661	BE325661
14	548	68.1	726	167	BE388478	BE388478
15	545	67.7	582	167	BE394469	BE394469
16	545	67.7	751	165	BE300398	BE300398
17	542	67.3	741	105	AL520417	AL520417
18	540	67.1	565	115	AW409904	AW409904
19	539	67.0	760	154	BE482283	BE482283
20	531	66.0	842	140	BE796516	BE796516
21	526	65.3	889	172	BE975906	BE975906
22	514	63.9	553	110	AV753592	AV753592
23	514	63.9	1040	173	BE105407	BE105407
24	513	63.7	962	172	BE026410	BE026410
25	511	63.5	954	174	BE167833	BE167833
26	502	62.4	1029	139	BE736544	BE736544
27	501	62.2	836	140	BE795522	BE795522
28	497	61.7	560	114	AW328549	AW328549
29	493	61.2	716	169	BE792903	BE792903
30	491	61.0	713	154	BE470793	BE470793
31	491	61.0	730	152	BE340392	BE340392
32	491	61.0	843	169	BE794134	BE794134
33	486	60.4	845	137	BE562997	BE562997
34	484	60.1	521	19	AI379313	AI379313
35	482	59.9	874	175	BE251461	BE251461
36	477	59.3	823	32	AV705178	AV705178
37	475	59.0	729	106	AL558439	AL558439
38	472	58.6	604	32	AV726915	AV726915
39	466	57.9	677	165	BE281236	BE281236
40	466	57.9	757	140	BE781787	BE781787
41	465	57.8	794	174	BE113765	BE113765
42	464	57.6	611	103	AL934480	AL934480
43	462	57.4	665	141	BE893687	BE893687
44	461	57.3	852	141	BE902190	BE902190
45	457	56.8	884	139	BE737217	BE737217

ALIGNMENTS

RESULT 1  
LOCUS AL541916 918 bp mRNA EST 16-FEB-2001  
DEFINITION AL541916 LTI\_FL002\_PL1 Homo sapiens cDNA clone CS0DE007YL04 5 prime  
/ mRNA sequence.  
ACCESSION AL541916  
VERSION AL541916.1 GI:12873445  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 918)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope

```

|||||
Db 722 AATATCATCTCACAGCTGAGGCCTGGCCAGATTACTTCTCCACCCCTCCCTACT 781
Qy 721 cacctggccagccctgtgctgggcccctgcaattggaatgttggccaga 769
|||||
Db 782 CACCTGCCAGCCCTGCTGCTGGGCCCTTCGAATTGGAATGTTGCCAGA 830
|||||

RESULT 2
AV650400 677 bp mRNA EST 07-SEP-2000
LOCUS AV650400 GLC Homo sapiens cDNA clone GLCED08 3', mRNA sequence.
DEFINITION AV650400
ACCESSION AV650400
VERSION AV650400.1 GI:9871414
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 677)
AUTHORS Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H.,
, Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J.,
Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu
,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.
TITLE Homo sapiens cDNA clone
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1..677
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GLCED08"
/clone_lib="GLC"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 136 a 158 c 248 g 132 t 3 others
ORIGIN
|||||
Query Match 79.6%; Score 641; DB 32; Length 677;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gccaggagggcgagtggaagtggccgtggggcggtatggactagctgcgtgtgcc 60
Db 8 gccaggagggcgagtggaagtggccgtggggcggtatggactagctgcgtgtgcc 67
|||||
Qy 61 ctgagacgtctcagggcggtatatactctgctggggcggtgcgtgcgtgcgtgcgcg 120
Db 68 CTGAGACGCTCAGCGGGCTATATCTCGTGGTGGGCGCGGTCTAGTCTCGGCGCAGC 127
|||||
Qy 121 qcagcaagacggtgcagtgagagagagtggtctctggcggtgtcccgagttcagcaga 180
Db 128 GCAGCAAGACGGGTGCAGTGAAGGAGAGTGGCGGTCTGCGCGGGGTCCCGACTTCAGCAGA 187
|||||
Qy 181 gcgcgtgcagccatggcccaatcaagtggagatgccatccacagcagtgagggtttt 240
Db 188 GCCCTGCACCATGGCCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTT 247
|||||
Qy 241 gaaggggagccaggggaacaaggtgaacctggcagagctgttcaagggaaggggtgtg 300
Db 248 GAAGGGGAGCCAGGGAACAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGAAGGGGTGTG 307

```

```

Qy 301 ctgtttgagttctctggggcccttcaccctgctgatttccaagacacacacctgccaggttt 360
Db 308 CTGTTTGAGTTCCTGGGGCCTTCACCCCTGGATGTTTCCAGACACACCTGCGCCAGGTTT 367
|||||
Qy 361 gtgagcagggctgaggtctctgaaagccaagggagtcacaggtggtggtcctgtctgagtt 420
Db 368 GTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGAGTCCAGGTGGTGGCCCTGCTGAGTGT 427
|||||
Qy 421 aatgatgccttggactggcagtggtggcgagcccaagggcggaagggcaaggttcgg 480
Db 428 AATGATGCTTTGTGACTGGCGAGTGGGGCGGAGCCCAAGGCGGAAGGTTTCGG 487
|||||
Qy 481 ctctgggtgatcccaactggggccctttgggaaggagacagactattactagatgatcgc 540
Db 488 CTCCTGGCTGATCCCACTGGGSCCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCG 547
|||||
Qy 541 ctggtgtccattttgggaatcgacgtctcaagaggttccatggttccatggtgtacagatggc 600
Db 548 CTGGTGTCCATCTTTGGGAATCGAGCTCTCAAGAGGTTTCTCCATGTTGTTACAGATGCG 607
|||||
Qy 601 atagtgaagccctggaatgtggaaccagatgacacagcct 641
Db 608 ATAGTGAAGGCCCTGAATGTGGAACCAACAGATGTCACAGGCCT 648
|||||

RESULT 3
BE784530 790 bp mRNA EST 20-OCT-2000
LOCUS BE784530 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3876972 5',
DEFINITION 601474153F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3876972 5',
mRNA sequence.
ACCESSION BE784530
VERSION BE784530.1 GI:10205815
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 790)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution Information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIAM9638 row: j column: 13
High quality sequence stop: 733.

FEATURES
source
1..790
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3876972"
/clone_lib="NIH_MGC_68"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.8 kb. Library constructed by Life
Technologies."

BASE COUNT 161 a 193 c 273 g 163 t
ORIGIN
|||||
Query Match 78.6%; Score 633; DB 140; Length 790;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 gagtggaaagtggcggtgggactagctggtggtgcgcctgagacgctc 71

```







cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L1CM1208 row: c column: 08  
 High quality sequence stop: 751.  
 Location/Qualifiers  
 1. .1064  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4336063"  
 /clone\_lib="NIH\_MGC\_48"  
 /tissue\_type="primary B-cells from tonsils (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: B-cells; Vector: pOTB7; Site\_1: XhoI;  
 Site\_2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCAGGAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by Ling  
 Hong in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."  
 274 a 221 c 364 g 205 t

BASE COUNT  
 ORIGIN

Query Match 72.7%; Score 585; DB 172; Length 1064;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-291;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ggagcgagtggaagtgccgtggcggtatggactagctggcggtgtgcgcctga 64  
 DB 2 GGAGCGGAGTGAAGTGGCGGTGGGGGTATGGGACTAGCTGGCGCTGTGCGCCCTGA 61

QY 65 gacgtcagcggtatatactctcgtggcggtggcggtgcagtcgtcgcgcagcgag 124  
 DB 62 GACGCTCAGCGGGTATATACCTCGTGGTGGCGGGTGGCGGTGCTGCTGGCGGCGAG 121

QY 125 caagacgtgcagtgagagagtgggcgtctgcggggtccgcagtttcagcagacgg 184  
 DB 122 CAAGCGGTGAGTGAAGGAGTGGCGCTGCGGGGTCCGCGTTCAGCAGAGCGG 181

QY 185 ctgcagccatggcccaatcaagtgaggatgcatccacagcagtgaggtgtttgaag 244  
 DB 182 CTGACGCATGGCCCAATCAAGTGGGAGATGCCATCCACGAGTGGAGGTGTTGAAG 241

QY 245 ggagccaggaacaagtggaacctggcagagctgttcaaggcggaagaggtgtgtgt 304  
 DB 242 GGGAGCCAGGGAACAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGAGGTTGCTGT 301

QY 305 ttgagttcctggggtcacccttgatgttccaaagacacacctgccaggtttgtg 364  
 DB 302 TTGAGTTCTGGGGCTTACCCCTCGATGTTCCAGACACACACTGCCAGGGTTGTGG 361

QY 365 agcaggtcaggtcctgaagccaaaggagttccaggtgtgtggcctgtcagtgattaatg 424  
 DB 362 AGCAGGCTGAGGCTCTCAAGGCCAAGGAGTCCAGGTGGTGGCTGTCTGAGTGTAAATG 421

QY 425 atgcttctgactggcgagtgggcgagcccaagcggaagcggaaggttcaggtcc 484  
 DB 422 ATGCTTTGTGACTGGCGAGTGGGGCGAGCCCAAGAGGGGAAGGTTCCGGTCC 481

QY 485 tggctgacccactggggcctttgggaagagagacagactattactagatgattcgtg 544  
 DB 482 TGGCTGATCCACTGGGGCTTTGGGAGGAGAGACAGACTTATTAAGATGATTCGCTGG 541

QY 545 tgtccattttggaatcagcgtctcaagaggttctccatgtgtg 589  
 DB 542 TGTCCATCTTTGGGAATCGACGCTCTCAAGAGGTTCTCCATGGTGG 586

RESULT 9  
 BE410064

LOCUS  
 DEFINITION

BE410064 959 bp mRNA EST 21-JUL-2000  
 601300587F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:3630548 5',  
 mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE

BE410064  
 BE410064.1 GI:9346514  
 EST.  
 human.

ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 959)  
 NIH-MGC <http://mgc.nci.nih.gov/>.

AUTHORS  
 TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)

JOURNAL  
 COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: ATCC

FEATURES

Location/Qualifiers  
 1. .959  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3630548"  
 /clone\_lib="NIH\_MGC\_21"  
 /tissue\_type="choriocarcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: placenta; Vector: pOTB7; Site\_1: XhoI;  
 Site\_2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCAGGAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."  
 203 a 255 c 314 g 187 t

BASE COUNT  
 ORIGIN

Query Match 72.5%; Score 584; DB 167; Length 959;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-291;  
 Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 ggactagctggcgtgtgcgcctgagacgtcagcggtcagcggtatatactcgtggggcc 99  
 DB 128 GGACTAGCTGGCGTGTGGCCCTCAGAGCGTCAAGGCGGTGAGTGGCGCTGTGGC 187

QY 100 ggcgttcagttctgcggcagcgagcagcggtgcagtcagtcagtcagtcagtcagtc 159  
 DB 188 GCGGTCAGTCTCGCGCAGCGGCGGAGAGCGGTGAGTGGAGAGTGGCGCTGTGGC 247

QY 160 ggggtccgcagtttcagcagagccgctgcagccatgcccccaatcaaggtgggagatgcc 219  
 DB 248 GGGGTCCGCAGTTTTCAGCAGAGCGGCTGCAGCCATGCCCCCAATCAAGTGGGAGATGCC 307

QY 220 atcccagcagtgaggtgtgttgaaggagagcagcgaggaacaggtggaacgtggagagctg 279  
 DB 308 ATCCCAGCAGTGGAGGTGTTGAAGGGGAGCCAGGGAACAGGTGAACCTGGCGAGCTG 367

QY 280 ttoagggcagaagggtgtgtgtttggagttcctggggccttccccccttgatgttcc 339  
 DB 368 TTCAAGGGCAAGAGGGTGTGCTGTTGGAGTTCTCTGGGCGCTTCAAGCGGTGATGTTC 427

QY 340 aagacacacctgcccaggtttgttgagcaggtcaggtcctgaagggagaggtccag 399  
 |||||||





```
Db 534 CTTGGTCCACTCTTTGGGATCGACCTCTCAAGAGTTCTCCATGGTGTGACAGATG 593
QY 599 goatagtgagccctgaatgtgaacacagatggcacagggcctcacctgcagctggcac 658
Db 594 GCATAGTGAAGGCCCTGAATGTGAACACAGATGCACAGGCTCACCTGCAGCCTGGCAC 653
QY 659 ccaatatacatcacagctcagggcctggccagattactcctcacccctccctat 718
Db 654 CCAATATCATCTCACAGCTCTGAGGCCCTGGCCAGATTAATCTCCTCACCCCTCCCTAT 713
QY 719 ctcaactccagccctgctggtggccctgcaattggaattgtggccagattcttcaa 778
Db 714 CTCACCTGCCAGCCCTGCTGTGGGCCCTGCAATTGGAATGTGGCAGATTTCTGCAA 773
QY 779 taaacactgtggttgcgg 798
Db 774 TAAACACTGTGTGTTGGG 793

RESULT 13
BG325661 710 bp mRNA EST 27-FEB-2001
LOCUS 602424378F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4562443 5',
DEFINITION mRNA sequence.
ACCESSION BG325661
VERSION BG325661.1 GI:13132098
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 710)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM1274 row: k column: 20
High quality sequence stop: 678.
FEATURES
Location/Qualifiers
1..710
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOT7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```
BASE COUNT 174 a 167 c 222 g 144 t 3 others
ORIGIN
Query Match 68.3%; Score 550; DB 152; Length 710;
Best Local Similarity 100.0%; Pred. No. 1.7e-273;
Matches 550; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 157 ggcggggccgcagcttcacagagccgcgtgcagcatggcccaatcaagggtggagat 216
|||||
```

```
Db 2 GCGGGGGTCCGCAGTTTTCAGCAGAGCCGCTGCAGCCCATGGCCCCAATCAAGTGGGAGAT 61
QY 217 gccatccacagcagtgaggtgtttgaaggggagccagggaaacaagtgaaacctggcagag 276
Db 62 GCCATCCACAGCAGTGCAGGTGTTTGAAGGGGAGCCAGGAACAAGTGAACCTGGCAGAG 121
QY 277 ctgttcaagggaagaagggtgtgtgtttgaggttcttggggccttccaccctcgatgt 336
Db 122 CTGTTCAAAGGGCAAGAGGTTGCTGTTTGGAGTTTCTTGGGGCCTTACCCTTGGATGT 181
QY 337 tccaaacacacactgccagggtttgtgagcaggtgaggtcctgaagcccaaggagatc 396
Db 182 TCCAAACACACACTGCCAGGTTTGTGGAGAGGCTGAGGCTCTGAAGGCCCAAGGAGTC 241
QY 397 caggtgtggtcctgtcgtgaggttaataatgatgccttctgactggcggcagtgggccgagcc 456
Db 242 CAGGTGTGGCCTGTCTGAGTGTAAATGATGCCCTTGTGACTGGCGAGTGGGCCGAGCC 301
QY 457 cacaaggcgaaggcaaggttcggtcctcctggctgatacccaactggggcctttgggaaggag 516
Db 302 CACAAGCGGAGGCAAGGTTCCGCTTCCCTGGCTGATCCCACTGGGGCCTTTGGGAAGGAG 361
QY 517 acagacttattactagatgatcgtgtgtgtccatcttggaaatcgactcctcaagag 576
Db 362 ACAGACTTATTACTAGATGATTCGCTGGTGTCCATCTTTGGGAATCGAGCTCTCAAGAGG 421
QY 577 ttctccatggtgtacagatggcagatgagcaggtgaagccctgaatgtgaaaccagatggcaca 636
Db 422 TTCTCCATGTTGTACAGGATGCATAGTGAAGGCCCTGAATGTGAACCAACATGGCACA 481
QY 637 ggctcactgcagcctgcacccaataatcatctcagctctgagccttgggccagat 696
Db 482 GGCTCACCCTGCAGCCTGGCACCACCAATATCATCTCACAGCTCTGAGGCCCTGGCCAGAT 541
QY 697 tacttctcc 706
Db 542 TACTTCTCTCC 551

RESULT 14
BG388478 726 bp mRNA EST 21-JUL-2000
LOCUS 601287129F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3621583 5',
DEFINITION mRNA sequence.
ACCESSION BG388478
VERSION BG388478.1 GI:93333843
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 726)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM295 row: i column: 08
High quality sequence stop: 638.
FEATURES
Location/Qualifiers
1..726
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_44"
/tissue_type="endometrium, adenocarcinoma cell line"
```



Mon Oct 22 14:46:50 2001

us-09-486-167a-1.oligo.rst

Page 13

Search completed: October 22, 2001, 13:21:17  
Job time: 4142 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2001, 10:41:25 ; Search time 1294.87 seconds  
(without alignments)  
5876.688 Million cell updates/sec

Title: US-09-486-l67A-1

Perfect score: 805

Sequence: 1 gccaggagcgagtggaag.....ttgtggttgcggaaaaaaa 805

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.\*

1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_est3.\*  
4: gb\_est4.\*  
5: gb\_est5.\*  
6: gb\_est6.\*  
7: gb\_est7.\*  
8: gb\_est8.\*  
9: gb\_est9.\*  
10: gb\_est10.\*  
11: gb\_est11.\*  
12: gb\_est12.\*  
13: gb\_est13.\*  
14: gb\_est14.\*  
15: gb\_est15.\*  
16: gb\_est16.\*  
17: gb\_est17.\*  
18: gb\_est18.\*  
19: gb\_est19.\*  
20: gb\_est20.\*  
21: gb\_est21.\*  
22: gb\_est22.\*  
23: gb\_est23.\*  
24: gb\_est24.\*  
25: gb\_est33.\*  
26: gb\_est34.\*  
27: gb\_est35.\*  
28: gb\_est36.\*  
29: gb\_est37.\*  
30: gb\_est38.\*  
31: gb\_est39.\*  
32: gb\_est40.\*  
33: em\_estba.\*  
34: em\_estfun.\*  
35: em\_esthum1.\*  
36: em\_esthum2.\*  
37: em\_esthum3.\*  
38: em\_esthum4.\*  
39: em\_esthum5.\*  
40: em\_esthum6.\*  
41: em\_esthum7.\*  
42: em\_esthum8.\*  
43: em\_esthum9.\*

44: em\_esthum10.\*  
45: em\_esthum11.\*  
46: em\_esthum12.\*  
47: em\_esthum13.\*  
48: em\_esthum14.\*  
49: em\_esthum15.\*  
50: em\_esthum16.\*  
51: em\_esthum17.\*  
52: em\_esthum18.\*  
53: em\_esthum19.\*  
54: em\_esthum20.\*  
55: em\_esthum21.\*  
56: em\_esthum22.\*  
57: em\_esthum23.\*  
58: em\_esthum24.\*  
59: em\_esthum25.\*  
60: em\_esthum26.\*  
61: em\_esthum27.\*  
62: em\_esthum28.\*  
63: em\_estin1.\*  
64: em\_estin2.\*  
65: em\_estin3.\*  
66: em\_estin4.\*  
67: em\_estin5.\*  
68: em\_estomi.\*  
69: em\_estom2.\*  
70: em\_estov1.\*  
71: em\_estov2.\*  
72: em\_estpl1.\*  
73: em\_estpl2.\*  
74: em\_estpl3.\*  
75: em\_estpl4.\*  
76: em\_estpl5.\*  
77: em\_estpl6.\*  
78: em\_estpl7.\*  
79: em\_estpl8.\*  
80: em\_estpl9.\*  
81: em\_estpl10.\*  
82: em\_estro1.\*  
83: em\_estro2.\*  
84: em\_estro3.\*  
85: em\_estro4.\*  
86: em\_estro5.\*  
87: em\_estro6.\*  
88: em\_estro7.\*  
89: em\_estro8.\*  
90: em\_estro9.\*  
91: em\_estro10.\*  
92: em\_estro11.\*  
93: em\_estro12.\*  
94: em\_estro13.\*  
95: em\_estro14.\*  
96: em\_estro15.\*  
97: em\_estro16.\*  
98: em\_estro17.\*  
99: em\_estro18.\*  
100: em\_estro19.\*  
101: em\_estro20.\*  
102: gb\_est25.\*  
103: gb\_est26.\*  
104: gb\_est27.\*  
105: gb\_est28.\*  
106: gb\_est29.\*  
107: gb\_est30.\*  
108: gb\_est31.\*  
109: gb\_est32.\*  
110: gb\_est41.\*  
111: gb\_est42.\*  
112: gb\_est43.\*  
113: gb\_est44.\*  
114: gb\_est45.\*  
115: gb\_est46.\*  
116: gb\_est47.\*

```
117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*
190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
Location/Qualifiers  
1. 918  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DE007YL04"  
/clone\_lib="LTI\_FL002\_PL1"  
/lab\_host="DH10B"  
/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand  
cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-stranded cDNA was digested with Not I  
and cloned into the Not I and Eco RV sites of the  
pCMVSPORT 6 vector. Library was constructed by Life  
Technologies. Contact : Feng Liang Life Technologies, a  
division of Invitrogen 9800 Medical Center Drive Rockville  
Maryland 20850, USA Fax : (1) 301 610 8371 Email :  
fliang@lifetech.com URL :  
http://fulllength.invitrogen.com"

BASE COUNT 218 a 225 c 296 g 178 t 1 others  
ORIGIN

Query Match 99.2%; Score 798.6; DB 106; Length 918;  
Best Local Similarity 99.5%; Pred. No. 8.7e-197;  
Matches 801; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 qccagagcgagtggaagtgccgtggcggtatggactagctggcgtgcgc 60  
DB 62 GCCAGAGCGGAGTGGGAAGTGGCCGTGGGCGGATATGGGACTAGCTGGCTGCGCC 121  
QY 61 ctgagacgtcagcggtctatactcgtcgtggcggtatggactagctggcgtgcgc 120  
DB 122 CTGAGACGCTCAGCGGGCTATATCTCGTCGCTGGGCGGCTCAGTCTCGCGCAGCG 181  
QY 121 gcagcaagcgtgcagtgaaagagagtgccgtctggcggtccaggttcagcaga 180  
DB 182 GCAGCAAGACGCTGAGTGAAGAGAGTGGGCGGCTCTGGCGGGGTCCGAGTTTCAGCAGA 241  
QY 181 gcgcctgcagcagcagcccaatcaagtgaggagatgccatcccgagtgagtggtt 240  
DB 242 GCCGCTGCAGCCTAGCCGCCATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTT 301  
QY 241 gaaggggagcaggggaacaagtgaaacctggcagagctgttcaagggcaagaaggtgtg 300  
DB 302 GAAGGGGAGCCAGGGAACAAGCTGACCTGGCAGAGCTGTTCAAGGGCAAGAAGGGTGTG 361  
QY 301 ctgtttgagttcctggcgccctccacctgagtttccaaagacacacacccctgcagggtt 360  
DB 362 CTGTTTGGAGTTCTCTGGGCGCTTCACCCCTGGATGTTCCACACACACACCTGCCAGGGT 421  
QY 361 gtgagcagcgtgagcgtcctgaagccaagggagtcaggtggtggtcgtgagtggtt 420  
DB 422 GTGGACGCGCTGAGGCTCTGAAGGCCAAGGAGTTTCAGTGTGGCTGTCTGAGTGT 481  
QY 421 aatgagccttctgactggcagtgaggccagcccaaggggaagggaggttcgg 480  
DB 482 AATGATGCTCTTGTGACTGCGGAGTGGGCGGCGCCACCAAGGCGGAAGGTTCCG 541  
QY 481 ctctggtgagtcacctggcgcccttgggaagagagacagacttattactagatgctg 540  
DB 542 CTCTGGCTGATCCCACTGGGCGCTTTGGGAAGAGACAGACTTATTACTAGATGATTCG 601  
QY 541 ctggtgtccatctttgggaatcgacgtcctcaagaggttctcctatggtgtacaggtg 600  
DB 602 CTGGTGTCCATCTTTGGGAATCAGCTCACAAGAGTGTCTCCCATGTGGTACAGGATGGC 661  
QY 601 atagtgaagggccctgaatgtggaaccagatggaagggccctcagcagcctgcagcc 660  
DB 662 ATAGTGAAGGCCCTGAATGTGAACAGAGATGGCAGGCGCTACCTGCAGCGCTGGCACC 721  
QY 661 aatatactctcacagctctgagggccctgagggccagattactctctccacccctccatct 720

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	798.6	99.2	918	106	AL541916	AL541916 AL541916
2	773.4	96.1	818	106	AL535869	AL535869 AL535869
3	766.8	95.3	850	105	AL520418	AL520418 AL520418
4	715.2	88.8	831	154	EG498369	EG498369 602544160
5	711	88.3	825	155	EG545942	EG545942 602573237
6	705.2	87.6	741	105	AL520417	AL520417 AL520417
7	701.8	87.2	889	172	BF975906	BF975906 602246103
8	692.6	86.0	790	140	BE784530	BE784530 601474153
9	689.8	85.7	1064	172	BF976165	BF976165 602245026
10	674.4	83.8	703	175	EG287024	EG287024 602382319
11	673.4	83.7	760	154	EG482283	EG482283 602526614
12	669	83.1	959	167	BE410064	BE410064 601300587
13	668.6	83.1	840	172	EG026788	EG026788 602294255
14	667	82.9	677	32	AV650400	AV650400 AV650400
15	666.4	82.8	842	140	BE796516	BE796516 601589855
16	663.6	82.4	813	153	EG438079	EG438079 602490343
17	660.6	82.1	734	106	AL541915	AL541915 AL541915
18	652.2	81.0	751	165	BE300398	BE300398 600944516
19	650.6	80.8	779	139	BE729013	BE729013 601562279
20	649.8	80.7	909	139	BE735282	BE735282 601569342
21	647.2	80.4	954	174	EG167833	EG167833 602339987
22	645.6	80.2	794	139	BE733545	BE733545 601566089
23	644.4	80.0	713	154	EG470793	EG470793 602511837
24	642.8	79.9	710	152	EG325661	EG325661 602443378
25	638.2	79.3	836	140	BE795522	BE795522 601592505
26	633.2	78.7	851	172	BF975982	BF975982 602246188
27	632	78.5	1029	139	BE736544	BE736544 601306118
28	630.2	78.3	757	140	BE781787	BE781787 601470589
29	630.2	78.3	845	137	BE562997	BE562997 601336252
30	625	77.6	794	174	EG113765	EG113765 602284293
31	624.4	77.6	768	139	BE745244	BE745244 601574011
32	622.2	77.3	787	173	EG109369	EG109369 602280422
33	622	77.3	622	175	EG255777	EG255777 602368257
34	620.4	77.1	918	150	BF530246	BF530246 602071421
35	620.2	77.0	910	153	EG403357	EG403357 602419031
36	616.2	76.5	746	174	EG166688	EG166688 602339089
37	616.2	76.5	758	168	BF679207	BF679207 602153784
38	615.4	76.4	874	175	EG251461	EG251461 602360956
39	613.4	76.2	811	141	BE907261	BE907261 602312751
40	611.4	76.0	1040	173	EG105407	EG105407 602347038
41	610.2	75.8	911	174	EG119624	EG119624 60347038
42	609.6	75.7	898	140	BE780877	BE780877 601469492
43	609.4	75.7	665	141	BE893687	BE893687 60146488
44	607.8	75.5	942	175	EG253661	EG253661 602366708
45	606.8	75.4	858	152	EG337084	EG337084 602434357

ALIGNMENTS

RESULT 1  
AL541916  
LOCUS AL541916 918 bp mRNA EST 16-FEB-2001  
DEFINITION AL541916 LTI\_FL002\_PL1 Homo sapiens cDNA clone CS0DE007YL04 5 prime  
, mRNA sequence.  
ACCESSION AL541916  
VERSION AL541916.1 GI:12873445  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 918)  
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope

```

Db 722 AATATCATCTCACAGCTCTGAGCCCTGGCCAGATTACTTCTCCACCCCTCCCTATCT 781
QY 721 caccctccagccctgctgctgggcccctgaattggaattgtggcagatttcttcaata 780
Db 782 CACCTGCCAGCCCTGCTGTGGGGCCCTGCAATTGGAATGTTGGCCAGAATTCTGCAATA 841
QY 781 aacactgtggttgcggaataaaa 805
Db 842 AACACTTGTGTTGGCGCAAAAAA 866

RESULT 2
AL535869 818 bp mRNA EST 13-FEB-2001
LOCUS AL535869 LTI_FL013_FBrn1 Homo sapiens cDNA clone CS0DF013YH13 5
DEFINITION prime, mRNA sequence.
ACCESSION AL535869
VERSION AL535869.1 GI:12799362
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 818)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
1. 818
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DF013YH13"
/dev_stage="pooled tissue from post conception fetuses (20
week, 24 week and 26 week)"
/lab_host="DH10B"
/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by Life
Technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA Fax : (31) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 176 a 201 c 274 g 165 t 2 others
ORIGIN

Query Match 96.18; Score 773.4; DB 106; Length 818;
Best Local Similarity 99.4%; Pred. No. 2.9e-190;
Matches 796; Conservative 1; Mismatches 2; Gaps 2;

QY 6 gagcgagtgtaagtcggtggcggtatggactagctgctgcccctag 65
Db 1 GAGCGGAGTGAAGTGGCCGTGGGGGGGATGGACTAGCTGGCGTGGC-CCTGAG 59

QY 66 acgtcagcggtatatactctgctggtggcgccggtgctgctgcccagc 125
Db 60 ACGCTACGCGGCTATATCTCGTGGTGGCGCGCGGTGAGTCTGCGCGAGCGCAGC 119

QY 126 aagcgtgcagtaagagagtggtgctgctggcggttcgcagtttcagacagcgc 185
Db 120 AAGACGCTGACAGTAAGAGAGTGGCGCTCTGGCGGGGTCCGACAGTTTCACAGAGCGC 179

QY 186 tgcagccatggcccaatcaagtgaggagatgccatcccgagcagtgaggtgttgagg 245

```

```

Db 180 TGCAGCCANGGCCCCCAATCAAGGTGGGAGATGCCATCCACAGCAGTGGAGGTGTTGAAGG 239
QY 246 ggagccagggaacaagtgtaacctggcagagctgttcaagggaaggggtgtgtctgtt 305
Db 240 GGAGCCAGGGAACAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGAGGGTGTGCTGTT 299
QY 306 tggagttctctggggccttcaaccttgatgtttccaagacacacctgccagggttttgga 365
Db 300 TGGAGTTCCTGGGGCCCTTCAACCCCTGGATGTTTCCAAGACACACACTGCCAGGGTTTGGGA 359
QY 366 gcaggctgagctctgaaggccaaggagtcagggtgggtggtgctctcaagtattatga 425
Db 360 GCAGGCTGAGGCTCTGAAGGCCAAGGGAGTTCAGGTGGTGGCTGTCTGAGTGTTAATGA 419
QY 426 tgccttgtgactggcagtggtggccgagcccccacaaagcggaggaaggttcggtcct 485
Db 420 TGCCTTGTGACTGGCGAGTGGGGCGGAGCCACAAAGCGGAAGGCAAGSTTCGGGCTCCT 479
QY 486 ggcctatccactggggcctttgggaagagagacagacttattactagattcgctggt 545
Db 480 GGCTGATCCACTGGGGCCCTTTGGGAAGGAGAGACAGACTTATTACTAGATGATTCGCTGCT 539
QY 546 gtcca-tctttgggaatcagctctcaagaggttctccatggtgtggtacaggatgcatag 604
Db 540 GTCCACTCTTTGGGAATCGACGCTCAAGAGGTTCTCCATGTTGGTGTACAGGATGGCATAG 599
QY 605 tgaagccctgaattgtgaaccagatggcacaggcctcactgcagcctggcaccacaata 664
Db 600 TGAAGGCCCTGAATGTGAACACAGATGGCACAGGCTCACCTGCAGCCTTGGCACCAATA 659
QY 665 tcatctcagctctgagggcctggccagattcttccaccctccctctcacc 724
Db 660 TCATCTCAGAGCTTGAGGCCCTGGGCCAGATTACTTCTCCACCCCTCCCTATCTCACC 719
QY 725 tcccagcctgtgctggggcctgcaattggaattgtggcagattttctgcaataaaca 784
Db 720 TGCCAGCCCTGTGCTGGGGCCCTGCAATTGGAATGTTGGCCAGATTTCTGCAATAAACA 779
QY 785 ctgtggtttgcggaataaaa 805
Db 780 CTGTGGTGTGGCGCAAAAA 800

RESULT 3
AL520418 850 bp mRNA EST 13-FEB-2001
LOCUS AL520418 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB006YE21 5
DEFINITION prime, mRNA sequence.
ACCESSION AL520418
VERSION AL520418.1 GI:12783911
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 850)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
1. 850
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB006YE21"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA"

```













QY 361 gtggagcagctgagctctgaagccaaaggaggtccaggtggtgctgtgagttg 420  
 Db 409 GTGGAGCAGCTGAGGCTCTGAAGGCCAAAGGAGTCAGAGTGGTGGCCCTGTCTGAGTGT 468  
 QY 421 aatgatgctcttctgactgagtcgagtcgagccgagcccaagagcggaagcgagtcg 480  
 Db 469 AATGATGCTCTTGTGACTGCGAGTGGGCCGAGCCACAAAGCGCAAGGCTTCGG 528  
 QY 481 ctctggtgctgacccactggggtcttggg-aaggagacagacttattactagatgac 539  
 Db 529 CTCCTGGCTGATCCACTGGGCGCTTTGGGCAAGGAGACAGACTTATTACTAGATGATC 588  
 QY 540 gctggttcacatcttgggaatcagctctcaagaggttctccatggtgtacagatgg 599  
 Db 589 GCTGGTGTCCATCTTTGGGAATCGAGCTCTCAAGAGGTTCTCCATGTTGTACAGATGG 648  
 QY 600 catagtgaagccctgaatgtgaaccagatggcacaggcctcacctgagcctggcacc 659  
 Db 649 CATAGTGAAGCCCTGAATGTGAACCCAGATGACAGGCTCACCTGCAGCCTGGCA-C 707  
 QY 660 caatatcatctcagctctgagccctgggcccagagattacttctccacc 710  
 Db 708 CAATATCATCTCAG-TCGTAGGCGCTGGGCCAGATTACTTCTCCACC 757

RESULT 12  
 BE410064 959 bp mRNA EST 21-JUL-2000  
 LOCUS 601300587F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:3630548 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BE410064  
 VERSION BE410064.1 GI:9346514  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 959)  
 NIH-MGC <http://mgi.nci.nih.gov/>  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at: [image.llnl.gov](http://image.llnl.gov)  
 Plate: L1CM318 row: n column: 21  
 High quality sequence start: 35  
 High quality sequence stop: 792.  
 Location/Qualifiers  
 1..959  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3630548"  
 /clone\_lib="NIH\_MGC\_21"  
 /tissue\_type="choriocarcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: placenta; Vector: pOTB7; Site\_1: XhoI;  
 Site\_2: EcoRI; CDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCAGAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."  
 203 a 255 c 314 g 187 t

BASE COUNT  
 ORIGIN

Query Match 83.1%; Score 669; DB 167; Length 959;  
 Best Local Similarity 94.9%; Pred. No. 3.9e-163;  
 Matches 713; Conservative 0; Mismatches 35; Indels 3; Gaps 2;  
 QY 1 gccagagagcggagtggaagtggccgtggcggtatgggactagctggcgctgagcc 60  
 Db 89 GCCAGAGCGGAGTGGAGTGGCGGTGGCGGGATTTGGACTAGCTGGCGTGTGGCCC 148  
 QY 61 ctgagacactcagcgggtctatatactcgtggtgggcccggcggtcagctctgcgagcg 120  
 Db 149 CTGACACGCTACGCGGCTATATCTCTGCGTGGGCGCGGCTCAGTCTCGCGCAGCG 208  
 QY 121 gcagcaagacagtggtcagtggaaggagagtggtggtggtgggggtccgcagtttcagcaga 180  
 Db 209 GCAGCAAGACGGTGCAGTGAAGGAGTGGCGTCTGGCGGGGTCCGAGGTTTCAGCAGA 268  
 QY 181 gccgtgcagccatagcccccaatcaagtggtgagatgccatcccgagtgaggtgtt 240  
 Db 269 GCCGTGACGCTAGCCCAATCAAGGTGGGAGATGCCATCCACAGCAGTGGAGGTGTT 328  
 QY 241 gaaggagagccaggggaacaagtggaacctggcagagctgttcaagggcaagaaggtgtg 300  
 Db 329 GAAGGGAGCCAGGGAACAAGCTGAACCTGGCAGAGCTGTTCAAGGGCAAGAAGGTGTG 388  
 QY 301 ctgtttggagttcctggggcctccaccctgagatgttccaagacacacacgtccaggttt 360  
 Db 389 CTGTTTGGAGTTCTCTGGGCTTCACTCCCTGGATGTTCCAAAGACACACTGCGCAGGGTT 448  
 QY 361 gtggagcagctgagctctgaagccaaaggagtcaggtggtggtggtggtggtggtggt 420  
 Db 449 GTGGAGCAGCTGAGGCTCTGAAGGCCAAGGAGTCCAGGTGGTGGCTGTCTGAGTGT 508  
 QY 421 aatgatgctcttctgactggtggtggtggtggtggtggtggtggtggtggtggtggtggt 480  
 Db 509 AATGATGCTCTTGTGACTGCGAGTGGCGGCGGAGCCACAAAGCGCAAGGCTTCGG 568  
 QY 481 ctctggtgacccactggggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 540  
 Db 569 CTCTGGCTGATCCCACTGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCG 628  
 QY 541 ctggttccatctcttgggaatcgacgtctcaagaggttctccatggtggtggtggtggtggtggt 600  
 Db 629 CTGCTGTCATCTTTGGGAATCGACGCTCAAGAGGTTCTCCATGTTGTGACAGGATGC 688  
 QY 601 atagtgaagccctgaatgtggaaccagatggcacaggcctcacctgcagcctggcacc 660  
 Db 689 ATAGTGAAGGCCCTGAATGTGG-ACCAGATGGCACAGGCTCACCTGCGAGCTGGCAGCC 747  
 QY 661 aatcatctct--cacagctctgagccctgggcccagattacttctccaccctctccctat 718  
 Db 748 AATATCATCTTCAACAGTCTGAGGCGCTTGGGCCAGATTACTTCTCTACAGCCCATTCCTT 807  
 QY 719 ctccactgcccagccctgtgctggggccctg 749  
 Db 808 ATTTTCAACCTTGGCAAGGCTTGTTCATG 838

RESULT 13  
 BG026788 840 bp mRNA EST 24-JAN-2001  
 LOCUS 602294255F1 NIH\_MGC\_86 Homo sapiens cDNA clone IMAGE:4388940 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG026788  
 VERSION BG026788.1 GI:12414756  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 840)  
 NIH-MGC <http://mgi.nci.nih.gov/>  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)

601 atagtgaaggccctgaatgtggaaccagatggcacaggcctcacctgcagcctggacccc

121 gcagcaagacggtcagtgtaaggagatggcgctctgcggggctcgcgagtttcacgaca 180  
|||||  
128 GCAGCAAGACGCTCAGTGAAGGAGAGTGGCGCTCTGGCGGGGTCCGCAGTTTCACGACA 187  
|||||

QY 181 gccctcagccatggcccccaatcaagtgaggatgccatccacagtgagggtgtt 240  
 Db 188 GCCCTCAGCCATGGCCCCAATCAAGTGGAGATGCATCCAGCAGTGGTGT 247  
 QY 241 gaaggagcaggaacaagtgaaactcagcagctgttcaaggcaagaagggtg 300  
 Db 248 GAAGGGAGCCAGGACAAAGTGAACCTGGCAGAGCTGTTCAGGGCAAGGGGTG 307  
 QY 301 ctgtttgagttctcctgggcccctcacccctggatgttccaaagacacacccctccagggttt 360  
 Db 308 CTGTTGGAGTCTCTGGGGCTTCAACCTTGATGTTCCAGACACACCTCCAGGGTT 367  
 QY 361 gtgagcagctgagctctgaagccaaaggagtgccagtggtgctctagtggt 420  
 Db 368 GTGAGCAGCTGAGGCTCTGAAGCCCAAGGGAGTCCAGGTGTGGCTGTCTGAGTGT 427  
 QY 421 aatgatcctttgtactggcgagtgggcgagccacacagccacagcggaaggttcgg 480  
 Db 428 AATGATGCTTGTGACTGGCGAGTGGGGCGAGCCACAAAGGCGGAAGGCAAGGTTCGG 487  
 QY 481 ctctggtgatccactggggccttgggaaggagacagacttattactagatgattcg 540  
 Db 488 CTCCTGCTGATCCACTGGGGCTTGGGAAGGAGACAGACTTATTACTAGATGATCG 547  
 QY 541 ctggtgcatctttgggaatogacgtctcaagaggttctccatgggtacagatggc 600  
 Db 548 CTGCTGCTCATCTTTGGGAATCGACGCTCAAGAGGTCTCCATGGTGTACAGGATGGC 607  
 QY 601 atagtgaagccctgaatgtggaaccagatggcacagcctcacctcagcctggcacc 660  
 Db 608 ATAGTGAAGCCCTGAATGTGGAAACAGATGGCAGGCTTACCTGCAGNCTGGCACCC 667  
 QY 661 aatatcatc 669  
 Db 668 AATATCATC 676

## RESULT 15

BE796516  
 LOCUS 60158985F1 NIH\_MGC\_7 Homo sapiens cDNA clone IMAGE:3944244 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BE796516  
 VERSION BE796516.1 GI:10217701  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 842)  
 NIH-MGC <http://mgi.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapb-femail.nih.gov](mailto:cgapb-femail.nih.gov)  
 Tissue Procurement: DCTD/DTF  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 Cloning distribution by: Incyte Genomics, Inc.  
 found through the I.M.A.G.E. Consortium/LNL at: [image.llnl.gov](http://image.llnl.gov)  
 Plate: L1CM800 row: m column: 13  
 High quality sequence stop: 711.

FEATURES  
source

1. .842  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3944244"  
 /clone\_lib="NIH\_MGC\_7"  
 /tissue\_type="small cell carcinoma"  
 /cell\_line="MGC3"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: lung; Vector: pOTB7; Site\_1: XhoI; Site\_2:

ECORI; cDNA made by oligo-dT priming. Directionally  
 cloned into EORI/XhoI sites using the following 5'  
 adaptor: GGCACAG(G). Size-selected >500bp for average  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 193 a 199 c 271 g 179 t  
 ORIGIN

Query Match 82.8%; Score 566.4; DB 140; Length 842;  
 Best Local Similarity 99.6%; Pred. No. 1.8e-162;  
 Matches 689; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 114 ggcagcgcagcaagcagtgacagtgaagagagtgagcgtcgtgaggggtccgcagttt 173  
 Db 4 GGCAGCGCAGACAGACGTCAGTGAAGAGAGTGGCGTCTGGCGGGTCCGAGTTT 63  
 QY 174 cagcagagccgtcgcagccatggccccaatcaagtgaggatgccatccacagcagtga 233  
 Db 64 CAGCAGAGCCGCTCAGCCATGGCCCAATCAAGGTGGAGATGCCATCCAGCAGTGA 123  
 QY 234 ggtgttgaaggagcaggggaacaagtgaaacctggcagagctgttcaaggggcaaga 293  
 Db 124 GGTGTTGAAGGGAGGCA-GGAACAGGTGAACCTGCAGAGCTGTTCAGAGGCAAGAA 182  
 QY 294 ggggtgctgtttgagttcctgggcttcaccctgagtggttccaaagacacacctcc 353  
 Db 183 GGGTGTGCTGTTTGAGTTCCTGGGGCTTCACCCCTGGATGTTCGAAGACACACCTGCC 242  
 QY 354 aggttgttgagcagcgtgaggtctgaagccaaaggagtcaggtgggtggtctct 413  
 Db 243 AGGTTGTGGAGCAGGCTGAGGCTCTGAAGCCCAAGGAGTCCAGGTGGTGGCTGCT 302  
 QY 414 gagtgttaatgatccctttgtgactggcgagtgggggccgagccacaaaggcggaaggcaa 473  
 Db 303 GAGTGTTAATGATGCTTTGTGACTGGCGAGTGGGGCGGAGCCACAAAGGGGAAGGCAA 362  
 QY 474 gtttcgctcctgctaatccactgggcttgggaaggagacagacttattactaga 533  
 Db 363 GTTTCGGCTCTGGCTGATCCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGA 422  
 QY 534 tgattcgtgtgtccatctttgggaatcgacgtctcgaaggttctccatgggtgtaca 593  
 Db 423 TGATTGCTGTGTCCTCTTTGGGAATCGAGTCTCAAGAGGTCTCCATGGTGTACA 482  
 QY 594 ggatggcatagtgaagccctgaaatgtggaaccagatggcacagcctcactcagcct 653  
 Db 483 GGATGGCATAGTGAAGGCCCTGAATGTGGAACCAAGATGGCACAGGCCCTCACCTGCAGCT 542  
 QY 654 ggcacccaataatcctcaagctcagcctgggagcctgggagcagattactcctccaccctc 713  
 Db 543 GGCACCAATAATCATCTCACAGCTGTAGAGCCCTGGGGCGGAGATTAATTCCTCCACCCCTC 602  
 QY 714 cctatctcactcctggcagcctgtgctgggctcgaatgggaatgttggccagatttc 773  
 Db 603 CCTATCTCACCTGCCAGCCCTGTGCTGGGGCCCTGCAATTGGAATGTGG-CAGATTTC 661  
 QY 774 tgcataaacaactgtgtgttgcggaaaaaa 805  
 Db 662 TGCAATAAACACTTGTGTGTTTGGCGCAAAAAA 693

Search completed: October 22, 2001, 11:50:25  
 Job time: 4140 sec



OM of US-09-486-167A-2 to: EST:\* out\_format : pfs

Date: Oct 22, 2001 10:19 AM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

```
-MODE=framer-p2n.model -DEV=xlp
-Q/cn2_1/USPTO_spool/US09486167/runat_22102001_064206_25012/app_query.fasta_1.222
-DB=EST -QFMT=fastp -SUFFIX=1st -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LCOEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -GAPOP=6.000
-DELOP=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09486167 @CNL_1.2970 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1
```

Search information block:

Query: US-09-486-167A-2

Query length: 162

Database: EST\*

Database sequences: 10228115

Database length: 431459454

Search time (sec): 1294.880000

score\_list:

Sequence	Strd	Orig	ZScore	Escore	Len	Documentation
gb_est72:BE298055	+	830.00	1740.13	8.3e-88	604	! BE298055 601118369F1 NIH_MGC_17
gb_est98:BG355777	+	830.00	1739.83	8.6e-88	622	! BG255777 602368257F1 NIH_MGC_91
gb_est95:BF971772	+	830.00	1739.81	8.6e-88	623	! BF971772 602240049F1 NIH_MGC_46
gb_est98:BG287024	+	830.00	1738.55	1.0e-87	703	! BG287024 602382319F1 NIH_MGC_93
gb_est99:BG325661	+	830.00	1738.21	1.0e-87	710	! BG325661 602424378F1 NIH_MGC_14
gb_est74:BE388478	+	830.00	1738.21	1.1e-87	726	! BE388478 601287129F1 NIH_MGC_44
gb_est29:AL541916	+	830.00	1735.76	1.5e-87	918	! AL541916 601306118F1 NIH_MGC_33
gb_est76:BE736544	+	830.00	1734.57	1.7e-87	1029	! BE736544 601306118F1 NIH_MGC_33
gb_est46:AW09904	+	826.00	1732.37	2.2e-87	565	! AW09904 601306118F1 NIH_MGC_33
gb_est40:AV650400	+	813.00	1702.98	9.7e-86	677	! AV650400 601306118F1 NIH_MGC_33
gb_est29:AL535869	+	810.00	1694.66	2.8e-85	817	! AL535869 601306118F1 NIH_MGC_33
gb_est79:BE784530	+	802.00	1678.10	2.4e-84	790	! BE784530 601474133F1 NIH_MGC_68
gb_est72:BE300059	+	802.00	1675.55	3.3e-84	1009	! BE300059 600944516F1 NIH_MGC_68
gb_est80:BE893687	+	801.00	1677.78	2.5e-84	665	! BE893687 601436488F1 NIH_MGC_72
gb_est72:BE300398	+	801.00	1677.01	2.7e-84	716	! BE300398 602253206F1 NIH_MGC_84
gb_est91:BF669149	+	801.00	1676.01	3.1e-84	788	! BF669149 602120006F1 NIH_MGC_17
gb_est79:BE795522	+	801.00	1675.39	3.3e-84	836	! BE795522 601592505F1 NIH_MGC_7
gb_est79:BE796516	+	801.00	1675.32	3.4e-84	842	! BE796516 601589855F1 NIH_MGC_7
gb_est92:BF794134	+	801.00	1675.31	3.4e-84	843	! BF794134 602255506F1 NIH_MGC_85
gb_est76:BE562997	+	801.00	1675.28	3.4e-84	845	! BE562997 601336252F1 NIH_MGC_44
gb_est78:BE737217	+	799.00	1668.48	3.6e-84	884	! BE737217 601305314F1 NIH_MGC_39
gb_est28:BF791499	+	796.00	1666.17	1.1e-83	734	! BF791499 602251549F1 NIH_MGC_8
gb_est28:AL541915	+	796.00	1665.88	1.1e-83	757	! AL541915 601470589F1 NIH_MGC_67
gb_est79:BE781787	+	796.00	1665.85	1.1e-83	757	! BE781787 601470589F1 NIH_MGC_67
gb_est21:BG498369	+	795.00	1664.88	1.3e-83	831	! BG498369 60244160F1 NIH_MGC_60
gb_est29:AL558438	+	795.00	1665.64	1.2e-83	631	! AL558438 601300587F1 NIH_MGC_21
gb_est28:AL558439	+	794.00	1661.84	1.4e-83	729	! AL558439 601300587F1 NIH_MGC_21
gb_est28:AL558437	+	794.00	1661.84	1.4e-83	729	! AL558437 601300587F1 NIH_MGC_21
gb_est78:BE729013	+	792.00	1657.09	3.5e-83	779	! BE729013 601562279F1 NIH_MGC_20
gb_est97:BG113765	+	790.00	1653.12	5.8e-83	760	! BG113765 602256614F1 NIH_MGC_21
gb_est74:BE410064	+	789.00	1652.66	6.2e-83	794	! BE410064 601300587F1 NIH_MGC_21
gb_est95:BF976165	+	789.00	1648.57	5.7e-83	611	! BF976165 602245026F1 NIH_MGC_4
gb_est28:AL520418	+	786.00	1641.14	2.7e-82	1064	! AL520418 601300587F1 NIH_MGC_21
gb_est95:BF975906	+	783.00	1632.44	4.5e-82	850	! BF975906 602245026F1 NIH_MGC_4
gb_est40:AV705034	+	777.00	1626.37	1.8e-81	707	! AV705034 602246103F1 NIH_MGC_48
gb_est102:BG545942	+	770.00	1609.95	1.5e-80	825	! BG545942 602573237F1 NIH_MGC_77
gb_est78:BE745244	+	768.50	1607.53	2.0e-80	768	! BE745244 601574011F1 NIH_MGC_9
gb_est16:AT119223	+	767.00	1608.81	1.7e-80	501	! AT119223 ue94h08.y1 Sugano mous

gb\_est13:AA869971 + 767.00 1607.90 1.9e-80 547 ! AA869971 vql9h12.r1 Barstead  
gb\_hic:AK002383 + 767.00 1604.38 3.0e-80 766 ! AK002383 Mus musculus adult  
gb\_hic:AK003332 + 767.00 1604.38 3.0e-80 766 ! AK003332 Mus musculus 18 day  
gb\_est100:BG403357 + 766.50 1601.53 4.4e-80 910 ! BG403357 602419031F1 NIH\_MGC\_1

seq\_name: gb\_est72:BE298055

seq\_documentation\_block: 604 bp mRNA EST  
LOCUS BE298055 601118369F1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:3028312 5',  
DEFINITION mrna sequence.

ACCESSION BE298055

VERSION BE298055.1 GI:9181641

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgi.nci.nih.gov/

1 (bases 1 to 604)

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LCM86 row: i column: 17

High quality sequence stop: 601.

Location/Qualifiers

1..604

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:3028312"

/clone\_lib="NIH\_MGC\_17"

/tissue\_type="rhabdomyosarcoma"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: muscle; Vector: pOTB7; Site: 1: EcoRI;

Site: 2: XhoI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Size-selected >500bp

for average insert size 1.8kb. Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 123 a 158 c 194 g 129 t

ORIGIN

alignment\_scores:

Quality: 830.00 Length: 162

Ratio: 5.123 Gaps: 0

Percent Similarity: 100.00 Percent Identity: 100.000

alignment\_block:

US-09-486-167A-2 x BE298055 ..

Align seg 1/1 to: BE298055 from: 1 to: 604

1 MetalaProleLysValGlyAspAlaIleProAlaValGluValPheG1 17

|||||

38 ATGGCCCAATCAAGTGGGAGATGCCATCCACGAGTGGAGTGTGA 87

|||||

17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34

|||||

88 AGGGAGCCAGGAAACAGGTGAACCTGGCAGAGCTGTTCAGGGCAAGA 137

|||||

34 yGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50

|||||

138 AGGGTGTGCTGTTGGAGTTCTGGGGCTTCACCCCTGGATGTTCCAAG 187

|||||

```
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLeuLysAlaLysG1 67
188 ACACACCTGCCAGGGTTTGTGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 237
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
238 AGTCCAGGTGGTGGCTGCTGAGTGTAAATGATGCTTGTGACTGGCG 287
84 luTriPlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
288 AGTGGGCGGAGCCACACAGGCGGAAGCAAGTTCTGGCTCTCGGCTGAT 337
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
338 CCCACTGGGGCTTTGGGAGGAGACAGACTTATTACTAGATGATTCGCT 387
117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
388 GGTGTCACTTTGGGAATCGAGCTCTCAGAGGTTCTCCATGTTGTGATC 437
134 luAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
438 AGGATGCATAGTGAAGGCCCTCAATGTGGAACACAGATGCGACAGGCCTC 487
151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
488 ACCTGCAGCTGGCACCACCAATATCATCTCACAGCTC 523
seq_name: gb_est98:BG255777
seq_documentation_block: 622 bp mRNA EST 13-FEB-2001
LOCUS BG255777 602368257F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4476207 5',
DEFINITION mRNA sequence.
ACCESSION BG255777
VERSION BG255777.1 GI:12765593
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 622)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
DCTD/DTP
Tissue Procurement: DCTD/DTP
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10303 row: b column: 16
High quality sequence stop: 622.
Location/Qualifiers
1..622
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4476207"
/clone_lib="NIH_MGC_91"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-3T primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
126 a 153 c 216 g 127 t
BASE COUNT
ORIGIN
```

```
alignment_scores: Quality: 830.00 Length: 162
Ratio: 5.123 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-486-167A-2 x BG255777 ..
Align seg 1/1 to: BG255777 from: 1 to: 622
1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
109 ATGGCCCAATCAAGGTGGAGATGCCATCCAGCAGTGGAGGTGTGA 158
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
159 AGGGAGCCAGGGAACAAGCTGAACCTGGCAGAGCTGTTCAGGGCAGA 208
34 yGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
209 AGGTGTGCTGTTTGGAGTTCTCTGGGGCTTCACTCCCTGGATGTTCCA 258
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
259 ACACACCTGCCAGGGTTTGTGAGCAGGCTCTCAGAGGTTCTCCATGTT 308
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
309 AGTCCAGGTGGTGGCTGCTGAGTGTAAATGATGCTTGTGACTGGCG 358
84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
359 AGTGGGCGGAGCCACACAGGCGGAAGGTTCTGGCTCTCGGCTGTGAT 408
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
409 CCCACTGGGGCTTTGGGAGGAGACAGACTTATTACTAGATGATTCGCT 458
117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
459 GGTGTCACTTTGGGAATCGAGCTCTCAGAGGTTCTCCATGTTGTGATC 508
134 luAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
509 AGGATGCATAGTGAAGGCCCTCAATGTGGAACACAGATGCGACAGGCCTC 558
151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
559 ACCTGCAGCTGGCACCACCAATATCATCTCACAGCTC 594
seq_name: gb_est95:BF971772
seq_documentation_block: 623 bp mRNA EST 22-JAN-2001
LOCUS BF971772 602240049F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4328364 5',
DEFINITION mRNA sequence.
ACCESSION BF971772
VERSION BF971772.1 GI:12338987
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 623)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

```
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
```

DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LCM1188 row: b column: 13  
 High quality sequence stop: 619.

# FEATURES

source  
 1. 623  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4328364"  
 /clone\_lib="NIH\_MGC\_46"  
 /tissue\_type="leiomyosarcoma cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: uterus; Vector: pOMB7; Site:1: XhoI; Site:2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCCAGGAG(G). Size-selected >500bp for average insert size  
 1.8kb. Library constructed by Ling Hong in the laboratory  
 of Gerald M. Rubin (University of California, Berkeley)  
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
 II RT (Life Technologies). Note: this is a NIH\_MGC  
 Library."  
 BASE COUNT 127 a 157 c 205 g 133 t 1 others  
 ORIGIN

alignment\_scores:  
 Quality: 830.00 Length: 162  
 Ratio: 5.123 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-486-167A-2 x BF91772 ..  
 Align seg 1/1 to: BF91772 from: 1 to: 623

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17  
 |||||  
 56 ATGGCCCAATCAAGTGGAGATGCCATCCAGCAGTGGAGGTGTTGA 105  
 |||||  
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysL 34  
 |||||  
 106 AGGGAGCCAGGACAAAGTGAACCTGCAGAGCTGTCAAGGGCAAGA 155  
 |||||  
 34 ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
 |||||  
 156 AGGGTGTGCTGTTGGAGTCTCTGGGGCTTCACCCCTGGATGTCCAAG 205  
 |||||  
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67  
 |||||  
 206 ACACACCTGCCAGGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 255  
 |||||  
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84  
 |||||  
 256 AGTCAGGTGGTGGCTGTCTGAGTGTAAATGATGCCCTTTGTGACTGGCG 305  
 |||||  
 84 lutrPglyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100  
 |||||  
 306 AGTGGGGCCGACCCACAAAGCGGAAGCAAGGTTCCGCTCTCTGGCTGAT 355  
 |||||  
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117  
 |||||  
 356 CCCACTGGGGCTTTGGGAAGGAGACAGACTATTACTAGATGATTCGCT 405  
 |||||  
 117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134  
 |||||  
 406 GGTGTCCATCTTTGGGAATGCACGCTCTCAAGAGGTCTCCATGGTGGTAC 455  
 |||||  
 134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150  
 |||||  
 456 AGGATGGCATAGTGAAGCCCTGAATGTGGAAACCAGATGGCACAGGCCTC 505  
 |||||

151 ThrCysSerLeuAlaProAsnIleSerGlnLeu 162  
 |||||  
 506 ACCTGCAGCCTGGCACCACCAATATCATCTACAGCTC 541  
 |||||

seq\_name: gb\_est98:BG287024

seq\_documentation\_block:  
 LOCUS BG287024 703 bp mRNA EST 21-FEB-2001  
 DEFINITION 602382319F1 NIH\_MGC\_93 Homo sapiens cDNA clone IMAGE:4499716 5',  
 mRNA sequence.

ACCESSION BG287024  
 VERSION BG287024  
 KEYWORDS EST.  
 SOURCE human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 703)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgabbs-re@mail.nih.gov](mailto:cgabbs-re@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10364 row: f column: 05

High quality sequence stop: 702.

Location/Qualifiers

## FEATURES

Source  
 1. 703  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4499716"  
 /clone\_lib="NIH\_MGC\_93"  
 /tissue\_type="transitional cell papilloma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: bladder; Vector: pCMV-SPORT6; Site:1: NotI;  
 Site:2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.7 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."  
 BASE COUNT 142 a 177 c 234 g 150 t  
 ORIGIN

## alignment\_scores:

Quality: 830.00 Length: 162  
 Ratio: 5.123 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-486-167A-2 x BG287024 ..

Align seg 1/1 to: BG287024 from: 1 to: 703

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17  
 |||||  
 105 ATGGCCCAATCAAGTGGAGATGCCATCCAGCAGTGGAGGTGTTGA 154  
 |||||  
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyL 34  
 |||||  
 155 AGGGAGCCAGGGAACAAGGTGAACCTGGCAGAGCTGTCAAGGGCAAGA 204  
 |||||  
 34 ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
 |||||  
 205 AGGGTGTGCTGTTGGAGTTCCTGGGGCTTCACCCCTGGATGTCCAAG 254  
 |||||  
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67  
 |||||

```

255 ACACACCTCCAGGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 304
67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
|||||
305 AGTCAGGTGGTGGCTGTCTGAGTGTAAATGATGCTTTGACTGGCG 354
|||||
84 luTrpGlyArgAlaHisTysAlaGluGlyLysValArgLeuLeuAlaAsp 100
|||||
355 AGTGGGGCCGAGCCACAAAGGCGGAAGCAAGTTCCGGCTCCTGGCTGAT 404
|||||
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuLeuAspSerLe 117
|||||
405 CCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCGCT 454
|||||
117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
|||||
455 GGTGTCCATCTTTGGGAATCGAGTCTCAAGAGGTTCCTCCATGGTGTAC 504
|||||
134 luAspGlyLeuValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
|||||
505 AGGATGGCATGTGAAGGCCCTGAATGTGAACCAAGATGGCAGAGCCTC 554
|||||
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162
|||||
555 ACTGCAGCTGGCACCAATATCATCTCACAGCTC 590
|||||

```

seq\_name: gb\_est99:BG325661

```

seq_documentation_block: 710 bp mRNA EST 27-FEB-2001
LOCUS BG325661 602424378F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4562443 5',
DEFINITION mRNA sequence.
ACCESSION BG325661
VERSION BG325661.1 GI:13132098
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 710)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/BTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1274 row: k column: 20
High quality sequence stop: 678.
Location/Qualifiers
1..710
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4562443"
/clone_lib="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
adaptor: GGCAGCAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

```

BASE COUNT 174 a 167 c 222 g 144 t

```

alignment_scores:
Quality: 830.00 Length: 162
Ratio: 5.123 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-486-167A-2 x BG325661 ..
Align seg 1/1 to: BG325661 from: 1 to: 710

1 MetAlaProIleLysValGlyAspAlaIleProAlaValAlaValPheG 17
|||||
38 ATGGCCCCAATCAAGGTGGAGATGCATCCACGAGTGGAGGTGTGA 87
|||||
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLys 34
|||||
88 AGGGAGCCAGGGAACAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 137
|||||
34 ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
|||||
138 AGGGTGTGCTGTTGGAGTTCCTGGGGCTTCACCCCTGGATGTTCCAAG 187
|||||
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG 67
|||||
188 ACACACCTGCCAGGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGG 237
|||||
67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
|||||
238 AGTCCAGGTGGTGGCTGTCTGAGTGTAAATGATGCTTTGTGACTGGCG 287
|||||
84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
|||||
288 AGTGGGGCCGAGCCACAAAGGCGGAAGGTTCCGGCTCCTGGCTGAT 337
|||||
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
|||||
338 CCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCGCT 387
|||||
117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
|||||
388 GGTGTCCATCTTTGGGAATCGAGTCTCAAGAGGTTCCTCCATGGTGTAC 437
|||||
134 luAspGlyLeuValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
|||||
438 AGGATGGCATGTGAAGGCCCTGAATGTGAACCAAGATGGCAGAGCCTC 487
|||||
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162
|||||
488 ACCTGCAGCTGGCACCAATATCATCTCACAGCTC 523
|||||

```

seq\_name: gb\_est74:BE388478

```

seq_documentation_block: 726 bp mRNA EST 21-JUL-2000
LOCUS BE388478 601287129F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3621583 5',
DEFINITION mRNA sequence.
ACCESSION BE388478
VERSION BE388478.1 GI:9333843
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 726)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```

DNA Sequencing by: Inocyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov

Plate: LCM295 row: 1 column: 08  
 High quality sequence stop: 638.

# FEATURES

source

1..726  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3621583"  
 /clone\_lib="NH\_MGC\_44"  
 /cissue\_type="endometrium, adenocarcinoma cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: uterus; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCAGGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 161 a 171 c 233 g 161 t

ORIGIN

## alignment\_scores:

Quality: 830.00 Length: 162  
 Ratio: 5.123 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-486-167A-2 x BE388478 ..

Align seg 1/1 to: BE388478 from: 1 to: 726

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17  
 |||||  
 50 ATGGCCCCAATCAAGGTGGAGATGCCATCCACGAGTGGAGGTGTTGA 99  
 |||||  
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34  
 |||||  
 100 AGGGAGCCAGGGAACAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 149  
 |||||  
 34 ySGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
 |||||  
 150 AGGTGTGCTGTTGGAGTTCCTGGGGCCTTCACCCCTGGATGTTCCAAG 199  
 |||||  
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysG1 67  
 |||||  
 200 ACACCTGCCAGGGTTTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 249  
 |||||  
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84  
 |||||  
 250 AGTCCAGGTGGTGGCTGTCTGAGTGTATGATGCTTGTGACGTGGCG 299  
 |||||  
 84 luTrpGlyArgAlaHisLysAlaGluGluLysValArgLeuLeuAlaAsp 100  
 |||||  
 300 AGTGGGCCAGCCACAGAGCGGAAGGTTCCGGCTCTCGGCTGCTGAT 349  
 |||||  
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117  
 |||||  
 350 CCCACTGGGGCCCTTGGGAGGAGACAGACTTATTACTAGATGATTCGCT 399  
 |||||  
 117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134  
 |||||  
 400 GGTGTCCATCTTGGGAATCGAGCTCTCAAGAGTTCTCCATGGTGGTAC 449  
 |||||  
 134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150  
 |||||  
 450 AGGATGGCATAGTGAAGGGCCCTGAATGTGGAACACAGATGGCAGGCCCTC 499  
 |||||  
 151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162  
 |||||

500 ACCTGCAGCTGGCACCACCAATATCATCTCACAGCTC 535

seq\_name: gb\_est29:AL541916

seq\_documentation\_block:

LOCUS AL541916 918 bp mRNA EST 16-FEB-2001  
 DEFINITION AL541916 LTI\_FLO02\_PL1 Homo sapiens cDNA clone CS0DE007YL04 5 prime  
 , mRNA sequence.

ACCESSION AL541916

VERSION AL541916.1 GI:12873445

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 918)

AUTHORS Li,W.B., Gruber,C., Jesse,J., Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

## FEATURES

source

1..918  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DE007YL04"  
 /clone\_lib="LTI\_FLO02\_PL1"  
 /lab\_host="DH10B"  
 /note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand  
 cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-stranded cDNA was digested with Not I  
 and cloned into the Not I and Eco RV sites of the  
 pCMVSPORT 6 vector. Library was constructed by Life  
 Technologies. Contact : feng liang Life Technologies, a  
 division of Invitrogen 9800 Medical Center Drive Rockville  
 , Maryland 20850, USA Fax : (1) 301 610 8371 Email :  
 fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com"

BASE COUNT 218 a 225 c 296 g 178 t 1 others  
 ORIGIN

## alignment\_scores:

Quality: 830.00 Length: 162  
 Ratio: 5.123 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-486-167A-2 x AL541916 ..

Align seg 1/1 to: AL541916 from: 1 to: 918

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17  
 |||||  
 254 ATGGCCCCAATCAAGGTGGAGATGCCATCCACGAGTGGAGGTGTTGA 303  
 |||||  
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34  
 |||||  
 304 AGGGAGCCAGGGAACAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 353  
 |||||  
 34 ySGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
 |||||  
 354 AGGTGTGCTGTTGGAGTTCCTGGGGCCTTCACCCCTGGATGTTCCAAG 403  
 |||||  
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67  
 |||||  
 404 ACACCTGCCAGGGTTTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 453  
 |||||  
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84  
 |||||  
 454 AGTTCAGGTGGTGGCTCTCTCAGTGTTAATGATGCCTTTGTGACTGGCG 503  
 |||||

```

84  luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
|||||
504  AGTGGGCGCCAGCCACAAAGCGGAAGGAGGTTCCGGCTCGGTGAT 553
|||||
101  ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
|||||
554  CCCACTGGGGCCCTTGGGAAGGAGACAGACTTATTACTAGATCGCT 603
|||||
117  uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
|||||
604  GGTGTCCACTTGTGGAACTGAGCTCTCAAGAGGTTCTCCATGGTGAT 653
|||||
134  lnaSpGlyIleValLysAlaLeuAsnValcluproAspGlyThrGlyLeu 150
|||||
654  AGGATGGCATAGTGAAGCCCTGAATGTGGACACAGATGGCAGGCCCT 703
|||||
151  ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
|||||
704  ACCTGAGCGCTGGCACCACCAATATCATCTCACAGCTC 739
|||||

seq_name: gb_est78:BE736544

seq_documentation_block:
LOCUS BE736544 1029 bp mRNA EST 15-SEP-2000
DEFINITION 601306118F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3640276 5',
mRNA sequence.
ACCESSION BE736544
VERSION BE736544.1 GI:10150536
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1029)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM344 row: d column: 05
High quality sequence stop: 658.
Location/Qualifiers
1. 1029
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3640276"
/clone_lib="NIH_MGC_39"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: pancreas; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies)."
```

```

BASE COUNT 299 a 234 c 321 g 175 t
ORIGIN

alignment_scores:
Quality: 830.00 Length: 162
Ratio: 5.123 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```

alignment_block:
US-09-486-167A-2 x BE736544
Align seg 1/1 to: BE736544 from: 1 to: 1029

1  MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheGl 17
|||||
51  ATGGCCCAATCAAGGTGGGAGATGCCATCCAGCAGTGGAGGTGTTGA 100
|||||
17  uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
|||||
101  AGGGGAGCCAGGACAAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 150
|||||
34  ySgLyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
|||||
151  AGGGTGTGCTGTTGGAGTTCCTGGGGCCCTCACCCCTGGATGTTCCAAG 200
|||||
51  ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysGl 67
|||||
201  ACACACCTGCCAGGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGG 250
|||||
67  yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
|||||
251  AGTCCAGGTGGTGGCCCTGTCTGAGTGTAAATGATGCCCTTTGTGACTGG 300
|||||
84  luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
|||||
301  AGTGGGCGCCAGCCACAAAGCGGAAGGAGGTTTCGGCTCGGTGAT 350
|||||
101  ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
|||||
351  CCCACTGGGGCCCTTGGGAAGGAGACAGACTTATTACTAGATCATCTCGT 400
|||||
117  uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
|||||
401  GGTGTCCACTTGTGGAACTGAGCTCTCAAGAGGTTCTCCATGGTGATC 450
|||||
134  lnaSpGlyIleValLysAlaLeuAsnValcluproAspGlyThrGlyLeu 150
|||||
451  AGGATGGCATAGTGAAGCCCTGAATGTGGACACAGATGGCAGGCCCTC 500
|||||
151  ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
|||||
501  ACCTGAGCGCTGGCACCACCAATATCATCTCACAGCTC 536
|||||

seq_name: gb_est46:AW409904

seq_documentation_block:
LOCUS AW409904 565 bp mRNA EST 29-JUN-2000
DEFINITION fh03all.x1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2960948 5',
mRNA sequence.
ACCESSION AW409904
VERSION AW409904.1 GI:6935445
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 565)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
```

Plate: LLCM55 row: B column: 21  
Seq primer: -21M13 forward primer (ABI).  
Location/Qualifiers

# FEATURES

source  
1..565  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2960948"  
/clone\_lib="NIH\_MGC\_17"  
/tissue\_type="rhabdomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: muscle; Vector: pORF7; Site\_1: EcoRI;  
Site\_2: XhoI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 119 a 133 c 197 g 116 t  
ORIGIN  
alignment\_scores:  
Quality: 826.00 Length: 161  
Ratio: 5.130 Gaps: 0  
Percent Similarity: 100.000 Percent identity: 100.000

## alignment\_block:

US-09-486-167A-2 x AW409904 ..

Align seg 1/1 to: AW409904 from: 1 to: 565

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17  
|||||  
82 ATGGCCCCAATCAAGTGGAGATCCATCCAGAGTGGAGGTGTGA 131  
|||||  
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysL 34  
|||||  
132 AGGGAGCCAGGAACAAGTGAACCTGGCAGAGCTTCAAGGCCAAGA 181  
|||||  
34 ySGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
|||||  
182 AGGTGTGCTGTTGGAGTCTCTGGGGCTTACCCCTGGATGTCCAAG 231  
|||||  
51 ThrHisLeuProGlyPheValGluAlaGluAlaLeuLysAlaLysG1 67  
|||||  
232 ACACACCTGCCAGGTTTGTGGAGCAGCTGAGGCTCTGAAGGCCAAGG 281  
|||||  
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84  
|||||  
282 AGTCCAGGTGGTGGCTGCTGAGTGTAAATGATGCCCTTTGTGACTGGCG 331  
|||||  
84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100  
|||||  
332 AGTGGGGCCGAGCCACAAAGCGGAAGCAAGGTTCGGCTCTGGCTGAT 381  
|||||  
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117  
|||||  
382 CCCACTGGGGCTTTGGGAAGAGACAGACTTATTACTAGATGATTCGCT 431  
|||||  
117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134  
|||||  
432 GGTGTCCATCTTTGGGAATGACGCTCTCAAGAGGTTCCTCCATGGTGTAC 481  
|||||  
134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150  
|||||  
482 AGGATGGCATAGTGAAGGCCCTGAATGGAAACAGATGGACAGGCGCTC 531  
|||||  
151 ThrCysSerLeuAlaProAsnIleIleSerGln 161  
|||||  
532 ACCTGCAGCCTGGCACCACCAATATCATCTCACAG 564  
|||||

seq\_name: gb\_est40:AV650400

## seq\_documentation\_block:

LOCUS AV650400 677 bp mRNA EST 07-SEP-2000  
DEFINITION AV650400 GLC Homo sapiens cDNA clone GLCED08 3', mRNA sequence.  
ACCESSION AV650400  
VERSION AV650400.1 GI:9871414  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 677)  
AUTHORS Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H.,  
Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J.,  
Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,W., Lu  
G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.  
Homo sapiens cDNA clone  
Unpublished (2000)  
COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.

## FEATURES

### source

1..677  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="GLCED08"  
/clone\_lib="GLC"  
/tissue\_type="corresponding non cancerous liver tissue"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2:  
XhoI"

BASE COUNT 136 a 158 c 248 g 132 t 3 others  
ORIGIN

## alignment\_scores:

Quality: 813.00 Length: 159  
Ratio: 5.113 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.371

## alignment\_block:

US-09-486-167A-2 x AV650400 ..

Align seg 1/1 to: AV650400 from: 1 to: 677

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17  
|||||  
200 ATGGCCCCAATCAAGTGGAGATGCCATCCAGCAGTGGAGGTGTGA 249  
|||||  
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysLysL 34  
|||||  
250 AGGGAGCCAGGACAAAGTGAACCTGGCAGAGCTTCAAGGGCAAGA 299  
|||||  
34 ySGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
|||||  
300 AGGTGTGCTGTTGGAGTCTCTGGGGCTTACCCCTGGATGTCCAAG 349  
|||||  
51 ThrHisLeuProGlyPheValGluAlaGluAlaLeuLysAlaLysG1 67  
|||||  
350 ACACACCTGCCAGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGG 399  
|||||  
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84  
|||||  
400 AGTCCAGGTGGTGGCTGCTGAGTGTAAATGATGCCCTTTGTGACTGGCG 449  
|||||  
84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100  
|||||

```

450 ACTGGGGCCGAGCCCAAGGGGAGGAGGTTCTGGCTCCTGGCTGAT 499
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
|||||
500 CCCACTGGGGCCTTTGGGAGGAGACAGACTTATTACTAGATCGCT 549
117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
|||||
550 GGTGTCCATCTTTGGGAATCGAGCTCTCAAGAGGTTCTCCATGGTG 599
134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGly 150
|||||
600 AGGATGGCATAGTGAAGGCCCTGAATGTGGAACCAAGATGGCACAG 649
151 ThrCysSerLeuAlaProAsnIleLe 159
|||||
650 ACCTGCAGNCTGGCACCCCAATATCATC 676

seq_name: gb_est29:AL535869
seq_documentation_block:
LOCUS AL535869 818 bp mRNA EST 13-FEB-2001
DEFINITION AL535869 LTI_FL013.FBrn1 Homo sapiens cDNA clone CS0DF013YH13 5
prime, mRNA sequence.
ACCESSION AL535869
VERSION AL535869.1 GI:12799362
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 818)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..818
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DF013YH13"
/clone_lib="LTI_FL013.FBrn1"
/dev_stage="pooled tissue from post conception fetuses (20
week, 24 week and 26 week)"
/lab_host="DH10B"
/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by Life
Technologies. Contact: Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
Maryland 20850, USA Fax: (1) 301 610 8371 Email:
fliang@lifetech.com URL:
http://fulllength.invitrogen.com"

BASE COUNT 176 a 201 c 274 g 165 t 2 others
ORIGIN

```

```

alignment_scores:
  Quality: 810.00 Length: 162
  Ratio: 5.031 Gaps: 1
Percent Similarity: 99.383 Percent Identity: 98.765

```

```

alignment_block:
US-09-486-167a-2 x AL535869 ..

```

```

Align seg 1/1 to: AL535869 from: 1 to: 818

```

```

2 AlaProIleLysValGlyAspAlaIleProAlaValGluValPheGluG1 18

```

```

190 GCCCAATCAAGGTGGAGATGCCATCCAGCAGTGGAGGTGTTTGAAGG 239
18 yGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysLys 35
|||||
240 GGAGCAGGGAAACAAGTGAACCTGGCAGAGCTGTTCAGGGCAAGAGG 289
35 yValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLysThr 51
|||||
290 GTGTGCTGTTTGGAGTTCTCTGGGGCCTTCACCCCTGGATGTTCCAA 339
52 HisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysGly 68
|||||
340 CACCTGCCAGGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCCAAGG 389
68 lGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyGlu 85
|||||
390 CCAGGTGGTGGCTGTCTGAGTGTATTATGATGCTTTGTGACTGGCG 439
85 rpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 101
|||||
440 GGGGCCGAGCCCAAGGCGAAGCAAGGTTCCGCTCCTGGCTGATCCC 489
102 ThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLeuVa 118
|||||
490 ACTGGGGCCTTTGGGAAAGGAGACAGACTTATTACTAGATGATTCG 539
118 lSer_IlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
|||||
540 GTCCACTCTTTGGGATCCACGCTCTCAAGAGGTTCTCCATGGTGTAC 589
135 AspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 151
|||||
590 GATGGCATAGTGAAGGCCCTGAATGTGGAACCAAGATGGCACAGGCT 639
151 rCysSerLeuAlaProAsnIleLeSerGlnLeu 162
|||||
640 CTGCAGCCTGGCACCCCAATATCATCTCACAGCTC 673

seq_name: gb_est79:BE784530
seq_documentation_block:
LOCUS BE784530 790 bp mRNA EST 20-OCT-2000
DEFINITION 601474153P1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3876972 5',
mRNA sequence.
ACCESSION BE784530
VERSION BE784530.1 GI:10205815
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 790)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsbbs-remail.nih.gov
Tissue Procurement: DCTD/DPH/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9638 row: j column: 13
High quality sequence stop: 733.
Location/Qualifiers
1..790
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3876972"
/clone_lib="NIH_MGC_68"

```

/tissue\_type="large cell carcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: lung; Vector: pCMV-SpORT6; Site: 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: oligo dT.  
 Average insert size 1.8 kb. Library constructed by Life  
 Technologies."  
 BASE COUNT 161 a 193 c 273 g 163 t  
 ORIGIN

alignment\_scores:  
 Quality: 802.00 Length: 162  
 Ratio: 4.951 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-486-167A-2 x BE784530 ..  
 Align seg 1/1 to: BE784530 from: 1 to: 790

1 MetAlaProLeuLysValGlyAspAlaIleProAlaValGluValPheG1 17  
 |||||  
 182 ATGGCCCCAATCAAGGTGGAGATGCCATCCAGCAGTGGAGTGTGGA 231  
 |||||  
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysL 34  
 |||||  
 232 AGGGAGCCAGGACCAAGTGAACCTGGCAGAGCTGTTCAGGGCAAGA 281  
 |||||  
 34 ySGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
 |||||  
 282 AGGGTGTGCTTTGGAGTCTCTGGGGCTTCACCCCTGGATGTTCCAAG 331  
 |||||  
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67  
 |||||  
 332 ACACACCTGCCAGGCTTTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 381  
 |||||  
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84  
 |||||  
 382 AGTCCAGGTGGTGGCTGTCTGAGTGTAAATGATGCCCTTTGTGACTGGCG 431  
 |||||  
 84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100  
 |||||  
 432 AGTGGGGCCGAGCCACAAAGCGGAAGCAAGGTTCGGCTCTCGCTGAT 481  
 |||||  
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117  
 |||||  
 482 CCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATCGCT 531  
 |||||  
 117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134  
 |||||  
 532 GGTGTCCATCTTTGGGATCCAGCTCTCAAGAGGTCTCCATGTTGGTAC 581  
 |||||  
 134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150  
 |||||  
 582 AGGATGGCATAGTAGAAGCCCTGAATGTGAACCATGTCACAGGCTC 631  
 |||||  
 151 ThrCysSerLeuAlaProAsnIleSerGlnLeu 162  
 |||||  
 632 AC.TGCACCTGGCACCACCAATATCATCTCACAGCTC 666

seq\_name: gb\_est72:BE300059

seq\_documentation\_block:  
 LOCUS BE300059 1009 bp mRNA EST 20-JUL-2000  
 DEFINITION 600944516T1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:2960859 3',  
 mRNA sequence.  
 ACCESSION BE300059  
 VERSION BE300059.1 GI:9183807  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

# REFERENCE

1 (bases 1 to 1009)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
 Plate: L1CM54 row: 0 column: 04  
 High quality sequence start: 20  
 High quality sequence stop: 724.  
 Location/Qualifiers

## FEATURES

source

1..1009

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2960859"  
 /clone\_lib="NIH\_MGC\_17"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: muscle; Vector: pOTB7; Site: 1: EcoRI;  
 Site 2: XhoI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."  
 BASE COUNT 239 a 307 c 247 g 216 t  
 ORIGIN

## alignment\_scores:

Quality: 802.00 Length: 162  
 Ratio: 4.951 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-486-167A-2 x BE300059/rev ..

Align seg 1/1 to reverse of: BE300059 from: 1 to: 1009

1 MetAlaProLeuLysValGlyAspAlaIleProAlaValGluValPheG1 17  
 |||||  
 555 ATGGCCCCAATCAAGGTGGAGATGCCATCCAGCAGTGGAGTGTGGA 506  
 |||||  
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysLysL 34  
 |||||  
 505 AGGGAGCCAGGCAACAAAGTGAACCTGGCAGAGCTGTTCAGGGCAAGA 456  
 |||||  
 34 ySGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
 |||||  
 455 AGGGTGTGCTTTGGAGTCTCTGGGGCTTCAC.CCTGATGTTCCAAG 407  
 |||||  
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67  
 |||||  
 406 ACACACCTGCCAGGCTTTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 357  
 |||||  
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84  
 |||||  
 356 AGTCCAGGTGGTGGCTGTCTGAGTGTAAATGATGCTTTGTGACTGGCG 307  
 |||||  
 84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100  
 |||||  
 306 AGTGGGGCCGAGCCCAAGCGGAAGCAAGGTTCGGCTCTCGCTGAT 257  
 |||||  
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117  
 |||||  
 256 CCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCGCT 207

117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134  
|||||  
206 GGTGTCCATCTTTGGGAATCGAGCTCTCAAGAGGTTCTCCATGGTGGTAC 157  
|||||  
134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150  
|||||  
156 AGATGGCATAGTGAAGGCCCTTGAATGTGGACCATGTCACAGAGCCCTC 107  
|||||  
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162  
|||||  
106 ACCTGCAGCTGGCACCACCAATATCATCTCACAGCTC 71  
|||||  
seq\_name: gb\_est80:BE893687  
seq\_documentation\_block: 665 bp mRNA EST 20-OCT-2000  
LOCUS BE893687 601436488F1 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:3921669 5',  
DEFINITION mRNA sequence.  
ACCESSION BE893687  
VERSION BE893687.1 GI:10355301  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 665)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTF  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM9754 row: p column: 22  
High quality sequence stop: 665.  
Location/Qualifiers  
1..665  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="3921669"  
/clone\_lib="NIH\_MGC\_72"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 2 kb. Library constructed by Life  
Technologies."  
BASE COUNT 138 a 165 c 216 g 146 t  
ORIGIN  
alignment\_scores:  
Quality: 801.00 Length: 162  
Ratio: 4.944 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-486-167A-2 x BE893687 ..  
Align seg 1/1 to: BE893687 from: 1 to: 665  
1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG 17  
|||||  
.66 ATGGCCCCAATCAAGGTGGAGATGCCATCCACAGTGGAGGTGTTGA 115  
|||||  
17 uGlyClnProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34  
|||||  
116 AGGGAGGCCAGG.ACAAGAGTGAACCTGGCAGAGCTGTTCAAGGCCAAGA 164  
|||||

34 ySglyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
|||||  
165 AGGGTGTCTGTTGGAGTTCTCTGGGCCCTTACCCCTGGATGTTCCAG 214  
|||||  
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG 67  
|||||  
215 ACACACCTGCCAGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGG 264  
|||||  
67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84  
|||||  
265 AGTCCAGGTGGTGGCCTCTGTAGTGTATATGATGCTTTGTGACGGCG 314  
|||||  
84 lUtrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100  
|||||  
315 AGTGGGGCCGAGCCACAGGCGGAAGGCTTCGGCTCTCGGTGAT 364  
|||||  
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117  
|||||  
365 CCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATCGCT 414  
|||||  
117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134  
|||||  
415 GGTGTCCATCTTTGGGAATCGAGCTCTCAAGAGGTTCTCCATGGTGGTAC 464  
|||||  
134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150  
|||||  
465 AGGATGGCATAGTGAAGGCCCTGAATGTGGAACCATGTCACAGAGCTC 514  
|||||  
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162  
|||||  
515 ACCTGCAGCTGGCACCACCAATATCATCTCACAGCTC 550  
|||||  
seq\_name: gb\_est92:BF792903  
seq\_documentation\_block: 716 bp mRNA EST 12-JAN-2001  
LOCUS BF792903 602253206F1 NIH\_MGC\_84 Homo sapiens cDNA clone IMAGE:4345488 5',  
DEFINITION mRNA sequence.  
ACCESSION BF792903  
VERSION BF792903.1 GI:12097888  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 716)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM9965 row: 1 column: 01  
High quality sequence stop: 623.  
Location/Qualifiers  
1..716  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="4345488"  
/clone\_lib="NIH\_MGC\_84"  
/tissue\_type="adrenal cortex carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site\_1:  
NotI; Site\_2: SalI; Cloned unidirectionally; oligo-dT  
primed. Average insert size 1.229 kb. Library enriched for  
full-length clones and constructed by Life Technologies."

BASE COUNT 168 a 174 c 223 g 151 t  
ORIGIN

Note: this is a NIH\_MGC Library."

alignment\_scores:  
Quality: 801.00 Length: 162  
Ratio: 4.944 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-486-167A-2 x BF792903 ..

Align seg 1/1 to: BF792903 from: 1 to: 716

```
1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
|||||
51 ATGCCCAATCAAGTGGGATGCCATCCAGCAGTGGAGGTGTTGA 100
|||||
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
|||||
101 AGGGAGCCAGG.AACAAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 149
|||||
34 ySglyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
|||||
150 AGGGTGTGCTTTGGAGTTCTTGGGGCCTTACCCCTGGATGTTCAAG 199
|||||
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
|||||
200 ACACACCTGCCAGGGTTTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 249
|||||
67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
|||||
250 AGTCCAGGTGGTGGCTGTCTGAGTGTATGATGCTTTGTGACTGGCG 299
|||||
84 lUtrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
|||||
300 AGTGGGGCCGAGCCACAAAGCGGAAGCAAGGTTCCGGCTCTCTGGCTGAT 349
|||||
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
|||||
350 CCCACTGGGGCCTTTGGAGGAGACAGACTTATTACTAGATGATTCGCT 399
|||||
117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
|||||
400 GGTGTCCATCTTTGGGAATCGAGCTCTCAAGAGGTTCCTCATGGTGTAC 449
|||||
134 lNaspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
|||||
450 AGGATGGCATAGTGAAGGCCCTGAAATGTGGAACCAACAGATGGCACAGGCCTC 499
|||||
151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
|||||
500 ACCTGCAGGCTGGCACCAATATCATCTCACAGCTC 535
```



OM of: US-09-486-167A-2 to: N\_Geneseq\_0601.\* out\_format : pfs  
 Date: Oct 22, 2001 10:44 AM  
 About: Results were produced by the GenCore software, version 4.5,  
 Copyright (c) 1993-2000 CompuGen Ltd.

## Command line parameters:

-MODEL=frame+2n.model -DEV=xlpl  
 -O=cpn2.1/USPTO.spool/US09486167/runat\_22102001\_064207\_25068/app\_query.fasta\_1.222  
 -DB=N\_Geneseq\_0601 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000  
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000  
 -GAPOP=4.500 -GAPEXT=0.050 -GAPOP=10.000 -GAPEXT=0.500  
 -GAPOP=6.000 -GAPEXT=7.000 -GAPOP=10.000 -GAPEXT=0.500  
 -DELOB=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62  
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
 -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
 -NORM-ext -MINLEN=0 -MAXLEN=2000000000  
 -USER=US09486167@cgn1\_1\_197 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLPXY  
 -WAIT -THREADS=1

## Search information block:

Query: US-09-486-167A-2

Query length: 162

Database: N\_Geneseq\_0601.\*

Database sequences: 730101

Database length: 313950809

Search time (sec): 109.540000

## score\_list:

Sequence	Strd	Orig	zScore	Escore	Len	Documentation
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	830.00	1773.91	7.7e-91	805		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	830.00	1771.67	1.0e-90	993		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	667.00	1427.68	1.5e-71	452		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	573.00	1221.39	4.1e-60	601		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	568.00	1210.53	1.9e-59	604		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	476.00	996.64	1.5e-47	2710		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	423.00	897.98	4.8e-42	553		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	364.00	772.55	4.6e-35	453		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	349.00	739.75	3.1e-33	469		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	315.00	667.34	3.4e-29	423		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	294.00	616.14	2.4e-26	727		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	289.00	614.28	3.0e-26	315		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	289.00	605.19	9.7e-26	737		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	288.00	603.98	1.1e-25	674		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	286.00	603.09	1.3e-25	489		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	286.00	599.52	2.0e-25	683		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	277.00	580.14	2.4e-24	678		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	270.00	563.27	2.1e-23	797		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	259.50	546.15	1.9e-22	473		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	250.50	500.33	6.7e-20	5556		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	250.50	456.00	2.0e-17	34980		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	247.50	514.73	1.1e-20	789		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	241.50	455.61	2.1e-17	58909		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	238.50	494.97	1.3e-19	811		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	236.00	489.58	2.7e-18	705		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	227.00	471.62	2.7e-18	705		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	219.50	471.28	1.0e-12	183012		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	218.50	455.72	2.1e-17	559		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	214.50	373.63	7.7e-13	534720		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	214.50	373.63	7.7e-13	536165		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	197.50	416.29	3.2e-15	320		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	194.00	404.05	1.6e-14	495		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	194.00	402.91	1.8e-14	551		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	194.00	402.85	1.8e-14	554		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	191.50	399.39	2.8e-14	462		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	191.50	395.19	4.2e-14	618		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	173.50	352.30	1.2e-11	991		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	128.50	260.38	1.6e-06	600		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	108.00	178.74	0.0e-01	19619		
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	101.00	201.58	0.0029	564		

seq\_name: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965

seq\_documentation\_block:

ID\_AA27965 standard; cDNA: 805 BP.

AC\_AA27965;

DT 08-JUN-1999 (first entry)

XX Human bronchoalveolar polypeptide, B18hum, coding sequence.

DE B18hum: bronchoalveolar protein; peroxisome-associated polypeptide;  
 lung injury; oxidative stress-related disorder; inflammatory disease;  
 cardiovascular disease; neurodegenerative disorder; allergic reaction;  
 amyotrophic lateral sclerosis; apoptosis; high bone mass syndrome;  
 osteopetrosis; osteoporosis-pseudoglioma syndrome; diagnosis;  
 Bardet-Biedl syndrome 1; therapy; ss.

OS Homo sapiens.

XX

PN WO9909054-A2.

XX 25-FEB-1999.

PD

XX

PF 20-AUG-1998; 98WO-BE00124.

XX

PR 20-AUG-1997; 97BE-0000692.

XX

PA (UYLO-) UNIV CATHOLIQUE LOUVAIN.

PA (UYMO-) UNIV MONS-HAINAUT.

XX

PI Bernard A, Falmagne P, Hermans C, Knoops B, Wattiez R;

DR WPI: 1999-180968/15.

XX

DR P-PSDB; AAY01079.

XX

XX New low molecular weight human broncho-alveolar polypeptide

PT useful for diagnosis and/or treatment of lung injuries and diseases,

PT and oxidative stress-related diseases and disorders, especially

PT inflammatory diseases

XX

PS Claim 8; Page 33-34; 45pp; English.

XX

CC This sequence encodes the human bronchoalveolar polypeptide, designated

CC B18hum of the invention. B18hum is a low molecular weight human,

CC peroxisome-associated broncho-alveolar polypeptide. A diagnostic device

CC featuring the polypeptide, polynucleotide and/or inhibitor is useful for

CC in vitro detection of lung injuries and diseases or oxidative

CC stress-related diseases and disorders, especially inflammatory diseases.

CC The device is also useful for monitoring such diseases or disorders in

CC patients or fluid samples. The polypeptide, polynucleotide and/or inhibitor

CC form pharmaceutical compositions useful in the prevention and/or

CC treatment of these diseases or disorders, especially specific

CC cardiovascular diseases (e.g. atherosclerosis), neurodegenerative

CC disorders (e.g. Alzheimer's disease and Parkinson's disease), amyotrophic

CC lateral sclerosis, apoptosis and inflammatory reactions, allergic

CC reactions (e.g. asthma, hay fever and eczema), high bone mass syndrome,

CC osteopetrosis, osteoporosis-pseudoglioma syndrome, and Bardet-Biedl

CC syndrome 1. The polypeptide is also useful as a specific marker of the

CC above diseases or disorders in a wide variety of tissues.

CC The discovery of the peroxisome-associated polypeptide

CC enables the development of diagnosis and treatment of peroxisomal

CC disorders.

XX

XX Sequence 805 BP; 163 A; 200 C; 276 G; 166 T; 0 other;

## alignment\_scores:

Quality: 830.00 Length: 162  
Ratio: 5.123 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-486-167A-2 x AAX27965 ..

Align seg 1/1 to: AAX27965 from: 1 to: 805

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17  
193 ATGGCCCAATCAAGTGGAGATGTCATCCAGCAGTGGAGGTGTTGA 242  
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34  
243 AGGGAGCCAGGGAACAAGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 292  
34 ySgLyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
293 AGGGTGTGCTGTTGGAGTTCCTGGGGCCTTCACCCCTGGATGTTCCAAG 342  
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67  
343 ACACACCTGCCAGGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 392  
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84  
393 AGTCCAGGTGGTGGCTGCTCAGTGTAAATGATGCTTTGTGACTGGCG 442  
84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100  
443 AGTGGGGCCGAGCCCAACAAGGGGAAGCAAGGTTCTGGCTCGCTGAT 492  
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117  
493 CCCATGGGGCTTTGGGAGAGACAGACTTATTACTAGATGATCGCT 542  
117 uValSerLeuPheGlyAsnArgLeuLysArgPheSerMetValValG 134  
543 GGTGTCCATCTTTGGGAATCGACGCTCTCAAGAGGTTCTCCATGGTGTAC 592  
134 InAspGlyLeuValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150  
593 AGGATGGCATAGTGAAGGCCCTTGAATGTGGAACCAAGATGGCAGGCCTC 642  
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162  
643 ACCTCGACCTGGCACCACCAATATCATCTCACAGCTC 678

seq\_name: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT: AAX56411

## seq\_documentation\_block:

ID AAX56411 standard; cDNA; 993 BP.

XX AC AAX56411;

XX AC AAX56411;

XX AC AAX56411;

XX AC AAX56411;

XX AC AAX56411;

XX AC AAX56411;

XX AC AAX56411;

XX AC AAX56411;

XX AC AAX56411;

XX AC AAX56411;

XX AC AAX56411;

XX AC AAX56411;

PF 14-OCT-1998; 98WO-US21730.  
XX  
PR 28-OCT-1997; 97US-0959004.  
XX  
PA (INCY-) INCYTE PHARM INC.  
XX  
PI Corley NC, Hillman JL, Lal P, Shah P, Yue H;  
XX  
DR WPI; 1999-326702/27.  
XX  
PT P-PSDB; AAY17388.

Human vesicle membrane-like proteins, useful for diagnosis, treatment and prevention of e.g. developmental disorders

Claim 7; Fig 1; 105pp; English.

The present sequence encodes human vesicle membrane protein-like protein 1 (VMP-1). VMP proteins, and their fragments, are used to treat or prevent developmental or vesicle-trafficking disorders, while their antagonists are used to treat or prevent immunological, reproductive or neoplastic disorders. Typical of many such disorders are anaemia; Cushing's syndrome; muscular dystrophy; cataract; cystic fibrosis; Addison's disease; ulcerative colitis; allergies; microbial infections; Addison's disease; cancer of breast, testis and prostate. VMP proteins are also used to raise specific antibodies (used to detect VMP in immunocassays (for diagnosis or monitoring), in competitive drug screens and to purify VMP from natural sources) and to screen for specific antagonists (potential therapeutic agents). VMP polynucleotides or their fragments, are used in hybridization assays to detect VMP in biological samples (e.g. for diagnosis, including detection of mutations and polymorphisms), optionally after amplification; to express recombinant VMP, including in vivo for gene therapy, and to map the corresponding genomic sequence.

SQ Sequence 993 BP; 176 A; 260 C; 358 G; 197 T; 2 other;

## alignment\_scores:

Quality: 830.00 Length: 162  
Ratio: 5.123 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-486-167A-2 x AAX56411 ..

Align seg 1/1 to: AAX56411 from: 1 to: 993

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17  
385 ATGGCCCAATCAAGTGGAGATGTCATCCAGCAGTGGAGGTGTTGA 434  
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34  
435 AGGGAGCCAGGGAACAAGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 484  
34 ySgLyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
485 AGGGTGTGCTGTTGGAGTTCCTGGGGCCTTCACCCCTGGATGTTCCAAG 534  
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67  
535 ACACACCTGCCAGGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 584  
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84  
585 AGTCCAGGTGGTGGCTCTGTGAGTGTAAATGATGCTTTGTGACTGGCG 634  
84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100  
635 AGTGGGGCCGAGCCCAACAAGGGGAAGCAAGGTTCTGGCTCGCTGAT 684  
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117  
|||||

685 CCCACTGGGGCCCTTTGGGAAGGAGACAGACTATTACTAGATGATTCCT 734  
 117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134  
 |||||  
 735 GTGTCCATCTTTGGGAATCGACGTCTCAAGAGGTTCTCCATGTGGGTAC 784  
 134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150  
 |||||  
 785 AGGATGGCATAGTGAAGGCCCTGGAATGTGAACACAGATGGCACAGCCCTC 834  
 151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162  
 835 ACCTGAGCTGGCACCCCAATATCATCTCACAGCTC 870

seq\_name: /SIDS1/gcgdata/geneseq/geneseq/NA2001.DAT:AAF92316

seq\_documentation\_block:

ID AAF92316 standard; cDNA; 452 BP.

AC AAF92316;

DT 15-MAY-2001 (first entry)

XX Bovine mammary tissue derived cDNA #29.

XX Bovine; mammary gland; cancer; tumour; angiogenesis; ss.

OS Bos taurus.

XX WO200114553-A1.

XX 01-MAR-2001.

XX 23-AUG-2000; 2000WO-NZ00166.

XX 23-AUG-1999; 99US-0150330.

XX (GENE-) GENESIS RES & DEV CORP LTD.

PA (NZPA-) NEW ZEALAND PASTORAL AGRIC RES INST LTD.

XX Havukkala IJ, Gleen M, Grigor MR, Molenaar AJ;

XX WPI; 2001-226619/23.

XX New polypeptides and polynucleotides encoding the polypeptides, which  
 PT are expressed in bovine mammary gland tissue, useful for stimulating  
 PT mammary gland growth or function, or inducing differentiation of milk  
 PT producing cells -

PS Claim 1; Page 49-50; 97pp; English.

XX The present invention relates to proteins derived from bovine  
 CC mammary gland cells. The invention is useful for stimulating  
 CC bovine mammary gland cell growth and function, inhibiting the  
 CC growth of various mammary gland cancer cells, inhibiting  
 CC angiogenesis and vascularization of tumours, or modulating  
 CC the growth of blood vessels in a mammal.

XX Sequence 452 BP; 100 A; 103 C; 145 G; 104 T; 0 other;

alignment\_scores:

Quality: 667.00 Length: 146

Ratio: 4.697 Gaps: 0

Percent Similarity: 97.260 Percent Identity: 86.986

alignment\_block:

US-09-486-167A-2 x AAF92316 ..

Align seg 1/1 to: AAF92316 from: 1 to: 452

1 MetaProIleLysValGlyAspAlaIleProAlaValGluValPheCl 17  
 |||||

14 ATGCCCCCGGATTAAAGTTGGAGATGCCATTCCATCGGTGGAGGATTATGA 63  
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34  
 |||||  
 64 AAGGACCCGGGCAACAAGGTGAACCTGGCAGAGCTGTTCAAGGCAAGA 113  
 34 ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
 |||||  
 114 AGGAGTGCCTGTTTGGCTCCCTGGGCGCTTACCCCTGGTGGTCCAAG 163  
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG 67  
 |||||  
 164 ACCACCTGCCAGGTTTCGTGGAGCAGGCTGATGCTCTGAAGGCCAAGG 213  
 67 yValGlnValAlaAlaCysLeuSerValAsnAspAlaPheValThrGly 84  
 |||||  
 214 GATCCAGTGGTGGCATGCTGACCGTTAATGATGCTTTGTAACCTCAAG 263  
 84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100  
 |||||  
 264 AGTGGCACGCGCCACAAGCAGAGGCAAGGTTCCGGCTCCTGGCAGAC 313  
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117  
 |||||  
 314 CCCAGTGGGACTTTGGGAAGGACAGATTTGTTACTTGATGATTCAC 363  
 117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134  
 |||||  
 364 GCTCTTCTCTTTGGGAATCACCAGCTGAAGAGGTTCTCCATGGTGATAG 413  
 134 InAspGlyIleValLysAlaLeuAsnValGluProAsp 146  
 |||||  
 414 AGGATGCATCGTCAATCCTGAACCTGGAGCCAGAT 451

seq\_name: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27966

seq\_documentation\_block:

ID AAX27966 standard; cDNA; 601 BP.

XX AAX27966;

XX 08-JUN-1999 (first entry)

XX Human bronchoalveolar polypeptide, B18hum, coding sequence.

XX B18hum; bronchoalveolar protein; peroxisome-associated polypeptide;  
 KW lung injury; oxidative stress-related disorder; inflammatory disease;  
 KW cardiovascular disease; neurodegenerative disorder; allergic reaction;  
 KW amyotrophic lateral sclerosis; apoptosis; high bone mass syndrome;  
 KW osteopetrosis; osteoporosis-pseudoglioma syndrome; diagnosis;  
 KW Bardet-Biedl syndrome 1; therapy; ss.

XX Homo sapiens.

OS

XX WO9909054-A2.

XX 25-FEB-1999.

XX 20-AUG-1998; 98WO-BE00124.

XX 20-AUG-1997; 97BE-0000692.

XX (UYLO-) UNIV CATHOLIQUE LOUVAIN.

XX (UYMO-) UNIV MONS-HAINAUT.

XX Bernard A, Falmagne P, Hermans C, Knoops B, Wattiez R;

XX WPI; 1999-180968/15.

XX New low molecular weight human broncho-alveolar polypeptide -  
 PT useful for diagnosis and/or treatment of lung injuries and diseases,  
 PT and oxidative stress-related diseases and disorders, especially  
 PT inflammatory diseases

XX PS Disclosure; Page 40; 45pp; English.

XX CC This sequence encodes the human bronchoalveolar polypeptide, designated

CC CC B18hum of the invention. B18hum is a low molecular weight human,

CC CC peroxisome-associated broncho-alveolar polypeptide. A diagnostic device

CC CC featuring the polypeptide, polynucleotide and/or inhibitor is useful for

CC CC in vitro detection of lung injuries and diseases or oxidative

CC CC stress-related diseases and disorders, especially inflammatory diseases.

CC CC The device is also useful for monitoring such diseases or disorders in

CC CC patients or fluid samples. The polypeptide, polynucleotide and inhibitor

CC CC form pharmaceutical compositions useful in the prevention and/or

CC CC treatment of these diseases or disorders, especially specific

CC CC cardio-vascular diseases (e.g. atherosclerosis), neurodegenerative

CC CC disorders (e.g. Alzheimer's disease and Parkinson's disease), amyotrophic

CC CC lateral sclerosis, apoptosis and inflammatory reactions, allergic

CC CC reactions (e.g. asthma, hay fever and eczema), high bone mass syndrome,

CC CC osteopetrosis, osteoporosis-pseudoglioma syndrome, and Bardet-Biedl

CC CC syndrome 1. The polypeptide is also useful as a specific marker of the

CC CC above diseases or disorders in a wide variety of tissues.

CC CC The discovery of the peroxisome-associated polypeptide

CC CC enables the development of diagnosis and treatment of peroxisomal

CC CC disorders.

XX CC

SQ Sequence 601 BP; 115 A; 162 C; 200 G; 124 T; 0 other;

alignment\_scores:

Quality: 573.00 Length: 113

Ratio: 5.071 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-486-167A-2 x AAX27966 ..

Align seg 1/1 to: AAX27966 from: 1 to: 601

50 LysThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaVal 66

|||||

173 AAGACACACCTCCAGGGTTTGTGGAGCAGGCTGAGGCTCTCAAGGCCAA 222

66 sclyValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrG 83

|||||

223 GGGAGTCCAGGGTGGGCTGTCTGAGTGTAAATGATGCTTTGTGACTG 272

83 lyGluTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAla 99

|||||

273 GCGAGTGGGGCGAGCCACACAGCGCGAAGGCAAGGTCGGCTCTGGCT 322

100 AspProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSe 116

|||||

323 GATCCCACTGGGGCCTTTGGGAAGGAGACAGACTTATTACTAGATGATTC 372

116 rleuValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValV 133

|||||

373 GCTGGTGTCATCTTTGGGAATPCGACGCTCAAGAGGTCTCCATGGTGG 422

133 alGlnAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGly 149

|||||

423 TACAGATGGCATATGAAGGCCCTGATGTGGACACAGATGGCACAGGC 472

150 LeuThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162

|||||

473 CTCACCTGCAGCCTGGCACCACCAATATCATCTCACAGCTC 511

seq\_name: /SIDS1/gcdata/geneseq/geneseq/NAL99.DAT: AAX27968

seq\_documentation\_block:

ID AAX27968 standard; cDNA; 604 BP.

XX AAX27968;

XX

DT 08-JUN-1999 (first entry)

XX DE Human bronchoalveolar polypeptide, B18hum, coding sequence.

XX KW

KW B18hum; bronchoalveolar protein; peroxisome-associated polypeptide;

KW lung injury; oxidative stress-related disorder; inflammatory disease;

KW cardiovascular disease; neurodegenerative disorder; allergic reaction;

KW amyotrophic lateral sclerosis; apoptosis; high bone mass syndrome;

KW osteopetrosis; osteoporosis-pseudoglioma syndrome; diagnosis;

KW Bardet-Biedl syndrome 1; therapy; ss.

XX OS

OS Homo sapiens.

XX PN

PN WO9909054-A2.

XX PD

PD 25-FEB-1999.

XX PF

PF 20-AUG-1998; 98WO-BE00124.

XX PR

PR 20-AUG-1997; 97BE-0000692.

XX PA

PA (UYLO-) UNIV CATHOLIQUE LOUVAIN.

XX PA

PA (UYMO-) UNIV MONS-HAINAUT.

XX PI

PI Bernard A, Falmagne P, Hermans C, Knoop B, Wattiez R;

XX DR

DR WPI; 1999-180968/15.

XX XX

XX New low molecular weight human broncho-alveolar polypeptide -

PT useful for diagnosis and/or treatment of lung injuries and diseases,

PT and oxidative stress-related diseases and disorders, especially

PT inflammatory diseases

XX XX

PS Disclosure; Page 40-41; 45pp; English.

XX CC

CC This sequence encodes the human bronchoalveolar polypeptide, designated

CC B18hum of the invention. B18hum is a low molecular weight human,

CC peroxisome-associated broncho-alveolar polypeptide. A diagnostic device

CC featuring the polypeptide, polynucleotide and/or inhibitor is useful for

CC in vitro detection of lung injuries and diseases or oxidative

CC stress-related diseases and disorders, especially inflammatory diseases.

CC The device is also useful for monitoring such diseases or disorders in

CC patients or fluid samples. The polypeptide, polynucleotide and inhibitor

CC form pharmaceutical compositions useful in the prevention and/or

CC treatment of these diseases or disorders, especially specific

CC cardio-vascular diseases (e.g. atherosclerosis), neurodegenerative

CC disorders (e.g. Alzheimer's disease and Parkinson's disease), amyotrophic

CC lateral sclerosis, apoptosis and inflammatory reactions, allergic

CC reactions (e.g. asthma, hay fever and eczema), high bone mass syndrome,

CC osteopetrosis, osteoporosis-pseudoglioma syndrome, and Bardet-Biedl

CC syndrome 1. The polypeptide is also useful as a specific marker of the

CC above diseases or disorders in a wide variety of tissues.

CC The discovery of the peroxisome-associated polypeptide

CC enables the development of diagnosis and treatment of peroxisomal

CC disorders.

XX CC

SQ Sequence 604 BP; 117 A; 160 C; 200 G; 127 T; 0 other;

alignment\_scores:

Quality: 568.00 Length: 162

Ratio: 4.814 Gaps: 1

Percent Similarity: 72.840 Percent Identity: 72.840

alignment\_block:

US-09-486-167A-2 x AAX27968 ..

Align seg 1/1 to: AAX27968 from: 1 to: 604

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17

|||||

161 ATGGCCCAATCAAGGTGGAGATGCCATCCAGCAGTGGAGGTGTTGA 210

17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34



```
51 ..... 51
1252 GGCAGAGAAATGCTTGAACCCAGGAGCGTAGGTTGCAGTGAGTGAGAT 1301
51 ..... 51
1302 CGTGCCTCTGCAGTCCAGCCTGGGTGAAGAGCGAGACTCGGTCTCAAAA 1351
51 ..... 51
1352 TGAATAAATAAAGAAAGAAACAAAGTAGAGACTGCAAAAGGAACAGTACC 1401
51 ..... 51
1402 GGGATGTTGGAGAAACATACATACATTAATCCACACCCCTGTTGG 1451
51 ..... 51
1452 TCCTGCTAAATGACAGGCACTGTGGAAGGTGCTTGGGACTCAGATAAATA 1501
51 ..... 51
1502 AGACAAAGATCTGCCATGGAAGTTACGCTCTGGACCAATAGGCATTAG 1551
51 ..... 51
1552 GTTTCATTCTGAGCTTCTAGTGCCCAAGGCAAAAGGAATAGAAATGGT 1601
51 ..... 51
1602 TTAGACAGCTCTCATTTGCTGATCAAAAGGTGTTGAGGCAGACACTGAG 1651
51 ..... 51
1652 AGGCCTGGAGATAAAGGTGGGTGGGGTGCAGATGCAGTTATCCCTTT 1701
52 .....His..LeuProGlyPheValGlu 58
1702 GCCGACCCCTTTGTCCTCCCTCAGACACACCTGCCAGGTTGTGGAG 1751
59 GlnAlaGluAlaLeuLysAlaLysGlyValGlnValValAlaCysLeuSe 75
1752 CAGGCTGAGGCTCTGAAGCCCAAGGAGTCCAGGTGGCTGCTGTCTGAG 1801
75 rValAsnAspAlaPheValThrGlyGluTrpGlyArgAlaHisLysAlaG 92
1802 TGTTAATGATGCCTTTGTGACTGGCGAGTGGGGCCGACCCACAAAGCGG 1851
92 luGlyLys..... 94
1852 AAGGCAAGGTGAGGTGAGGGGCTGCAGGAGTGCAGGACAGGTAGGATA 1901
95 .....ValArgLeuLeuAlaAspPr 101
1902 TTCTTTCTGTCACCTCTACTTCTCTGCAGGTTCCGCTCCTGGCTGATCC 1951
101 oThrGlyAlaPhe GlyLys..... 107
1952 CACTGGGCCCTTTGGGAAGGTGAGTGTCCCTGTGACCCACAGGGACAT 2001
107 ..... 107
2002 GSCGTGCGGGAGCAGTGGGGCCCTTGGCCTCTTCAAGGATTTCTGCAC 2051
108 .....GluThrAspLeuLeuLeuAspSerL 117
2052 ACTTTTCTCTGCTCTCTTGTAGGACAGACACTATTACTAGTATCGC 2101
117 euValSerIlePheGlyAsnArgArgLeu..Lys..... 127
2102 TGGTGTCCATCTTTGGGAATCGACCTCTTCAGAGGTTAAAGTGGAGAGTC 2151
```

```
127 ..... 127
2152 CTCTGTGGAGAAAGTCTCTCTGTGGGAGAGAGTCTCTGTGGGAGAGATC 2201
127 ..... 127
2202 CTCTGTGTGAGAGAGGTCTCTCTGTGGAAAGAGTCTCTGTGGGGGAGATGTG 2251
127 ..... 127
2252 TGGGAGAGAGTCTCTGTGTGGGAGAGTCTCTGTAGGGGAGAGTCTCTCTG 2301
127 ..... 127
2302 GGGAGAGAGTCTCTGTGTGGGGGAGAGTCTCTGTGGGAGAGTCTCTCTGT 2351
127 ..... 127
2352 GTGGAGAGAGTCTCTGTGTGGTGTGAGTCTCTGTGGGGGAGAGTCTCTCT 2401
127 ..... 127
2402 GTGGGGGAGTCTCTCTGTGGAGTCTCTTGTGGCCCTGGCTGTTCACTGC 2451
127 ..... 127
2452 CTGTCTCCATGCCAGCCTCCAAGCCAGCTGATGCAGCTGGCTGGGCC 2501
128 .....ArgPheSerMetValValGlnAspGlyIleValLysAl 140
2502 CCTCTTCCGGCAGGTTCTCCATGGTGTACAGGATGGCATAGTGAAGGC 2551
140 aLeuAsnValGluProAspGlyThrGlyLeuThrCysSerLeuAlaProA 157
2552 CCTGAATGTGAACACAGATGCACAGGCTCACCTGCAGCCTGGCACCCA 2601
157 snIleIleSerGlnLeu 162
2602 ATATCATCTCACAGCTC 2618
seq_name: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT:AAC95054
seq_documentation_block:
ID AAC95054 standard; cDNA; 553 BP.
XX
AC AAC95054;
XX
DT 19-FEB-2001 (first entry)
XX
DE Cat flea hindgut and Malpighian tubule (HMT) cDNA, SEQ ID NO:1549.
XX
KW Cat flea; hindgut and Malpighian tubule nucleic acid; HMT;
KW flea infestation; vaccine; antiparasitic; therapeutic target;
KW diagnosis; detection; ss.
XX
OS Ctenocephalides felis.
XX
PN WO200061621-A2.
XX
PD 19-OCT-2000.
XX
PF 07-APR-2000; 2000WO-US09437.
XX
PR 09-APR-1999; 99US-0128704.
XX
PA (HESK-) HESKA CORP.
XX
PI Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N;
XX
DR WPI; 2000-656323/63.
XX
PT Flea Malpighian tubule and head and nerve cord tissue derived nucleic
acids useful for the prevention, diagnosis and treatment of flea
```

PT infestations -

XX Claim 26; Page 756; 964pp; English.

XX The invention relates to novel cat flea (*Ctenocephalides felis*) nucleic acids which are expressed in hindgut and Malpighian tubule (HMT) tissue or head and nerve cord (HNC) tissue. The invention also relates to the encoded proteins. The invention additionally encompasses expression constructs, recombinant viruses and recombinant cells comprising the nucleic acids of the invention, recombinant production of the proteins, antibodies against the proteins, a method of identifying inhibitors of the proteins, and compositions comprising the inhibitors for administration to an animal. The nucleic acids, and the proteins they encode may be used in the prevention, treatment and diagnosis of diseases associated with flea infestations. For example, the nucleic acids may be used to produce an HMT or HNC protein according to standard recombinant DNA methodology by inserting the nucleic acids into a host cell and culturing the cell to express the protein. The HMT and HNC nucleic acids may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect and quantitate the presence of cat flea or other homologous nucleic acid sequences in samples. They may also be used to study the expression and function of the proteins and their role in metabolism. The HMT and HNC proteins may be used as antigens in the production of specific antibodies, and in assays to identify modulators (agonists and antagonists) of HMT and/or HNC protein expression and activity. The anti-HMT/HNC protein antibodies and antagonists may also be used to down-regulate protein expression and activity. The antibodies may also be used as diagnostic agents for detecting the presence of flea polypeptides in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The present sequence represents a cat flea HMT cDNA of the invention.

XX Sequence 553 BP; 180 A; 86 C; 123 G; 164 T; 0 other;

alignment\_scores:

Quality: 423.00 Length: 164  
Ratio: 3.439 Gaps: 3  
Percent Similarity: 75.000 Percent Identity: 55.488

alignment\_block:

US-09-486-167A-2 x AAC95054 ..

Align seg 1/1 to: AAC95054 from: 1 to: 553

1 MetAlaProLeuValGlyAspAlaIleProAlaValGluValPheG1 17  
24 ATGGCTCAGATTAAAGTAGGCGATAAAATTCGCTCGGTAGACTATTGA 73  
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34  
74 AGACATCCAGCTAACAAAGTGAATATGCAATCTCGCTGCGAGGAAAA 123  
34 ySgLyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
124 AAGTTGCTTATTTCGCTGTTCTGCGCCCTTCACTCCAGGATGTTCTAG 173  
51 ThrHisLeuProGlyPheValGluAlaGluAlaLeuLysAlaLysG1 67  
174 ACTCATGTACTGCTGCTATGTTGCAAGGCTGAAGAATTAAGAGAGTGG 223  
67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrClyG 84  
224 AATTGCTGAATCTTTGCGTCTCTGTAATGATFCCTTTGTTATGATGT 273  
84 luTrpGlyArgAlaHisLysAlaGluLysValArgLeuLeuAlaAsp 100  
274 CTTGGGTGAAGATCAGCAGCAAGCAATGAAAGTAAAGATGCTTGCAGAT 323  
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerL 117  
324 CCAAGTGGTACATTCAAAAAGAACTTGGATTGGAGTTG..... 363  
117 euValSerIlePheGlyAsnArgArgLeuLysArgPhe..... 129

364 .....ATTTGCGCCCTTAGGAGTTTACGTTCTTAAGAG 396  
130 ...SerMetValValGlnAspGlyIleValLysAlaLeuAsnValGluPr 145  
397 ATCTATGTTGATGACAAATGCTGATCAGAAATGTAATGTAGAACC 446  
145 oAspGlyThrGlyLeuThrCysSerLeuAlaProAsnIle 158  
447 AGATGGTCTTGGACTTTCCTGTTCTTAGCTGACAAACTT 485

seq\_name: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT:AAC04049

seq\_documentation\_block:

ID AAC04049 standard; cDNA; 453 BP.

XX AAC04049;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 4047.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
XX gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GBST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX P-PSDB; AAG04043.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX Claim 1; SEQ ID 4047; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from  
XX mRNAs encoding secreted proteins. An ORF has been identified within the  
XX sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs  
XX derived from 30 different tissues. EST sequences usually correspond  
XX mainly to the 3' untranslated region (UTR) of the mRNA because they are  
XX often obtained from oligo-dT primed cDNA libraries. Such ESTs are not  
XX well suited for isolating cDNA sequences derived from the 5' ends of  
XX mRNAs and even in those cases where longer cDNA sequences have been  
XX obtained, the full 5' UTR is rarely included. 5' ESTs are derived from  
XX mRNAs with intact 5' ends and can therefore be used to obtain full length  
XX cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
XX gene therapy and chromosome mapping procedures. They are used to obtain  
XX upstream regulatory sequences and to design expression and secretion  
XX vectors.

XX Sequence 453 BP; 82 A; 108 C; 180 G; 82 T; 1 other;

alignment\_scores:

Quality: 364.00 Length: 72  
Ratio: 5.127 Gaps: 0  
Percent Similarity: 98.611 Percent Identity: 98.611

alignment\_block:

US-09-486-167A-2 x AAC04049 ..

Align seg 1/1 to: AAC04049 from: 1 to: 453

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17  
 226 ATGCCCCCAATCAAGGTGGAGATGCCATCCAGCAGTGGAGGTGTTGA 275  
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34  
 276 AGGGAGCCAGGACACAGGTGACCTGGCAGACSTGTTCAAGGGCAGA 325  
 34 yGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
 336 AGGGTGTCTGTTGGAGTCTCTGGGGCTTCAACCTGGATGTTCCCAAG 375  
 51 ThrHisLeuProGlyPheValGluGlnAlaGluLeuLysAlaLysG1 67  
 376 ACACACCTGCCAGGGTTTGTGGAGCAGCTGAGGCTCTGAAGGCCAAGG 425  
 67 yValGlnValValAla 72  
 426 AGTCAGGTGGTGGCT 441

seq\_name: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AAx27967

seq\_documentation\_block:

ID\_AAX27967 standard; cDNA; 469 BP.

XX AAX27967;

08-JUN-1999 (first entry)

Human bronchoalveolar polypeptide, B18hum, coding sequence.

B18hum; bronchoalveolar protein; peroxisome-associated polypeptide;  
 lung injury; oxidative stress-related disorder; inflammatory disease;  
 cardiovascular disease; neurodegenerative disorder; allergic reaction;  
 amyotrophic lateral sclerosis; apoptosis; high bone mass syndrome;  
 osteopetrosis; osteoporosis-pseudoglioma syndrome; diagnosis;  
 Bardet-Biedl syndrome 1; therapy; ss.

XX Homo sapiens.

PN WO9909054-A2.

XX 25-FEB-1999.

XX 20-AUG-1998; 98WO-BE00124.

XX 20-AUG-1997; 97BE-0000692.

XX (UYLO-) UNIV CATHOLIQUE LOUVAIN.

PA (UYMO-) UNIV MONS-HAINAUT.

XX Bernard A, Falmagne P, Hermans C, Knoop B, Wattiez R;

XX WPI; 1999-180968/15.

XX New low molecular weight human broncho-alveolar polypeptide -  
 useful for diagnosis and/or treatment of lung injuries and diseases,  
 PT and oxidative stress-related diseases and disorders, especially  
 PT inflammatory diseases

XX Disclosure; Page 39; 45pp; English.

XX This sequence encodes the human bronchoalveolar polypeptide, designated  
 CC B18hum of the invention. B18hum is a low molecular weight human,  
 CC peroxisome-associated broncho-alveolar polypeptide. A diagnostic device  
 CC featuring the polypeptide, polynucleotide and/or inhibitor is useful for  
 CC in vitro detection of lung injuries and diseases or oxidative  
 CC stress-related diseases and disorders, especially inflammatory diseases.  
 CC The device is also useful for monitoring such diseases or disorders in  
 CC patients or fluid samples. The polypeptide, polynucleotide and inhibitor

CC form pharmaceutical compositions useful in the prevention and/or  
 CC treatment of these diseases or disorders, especially specific  
 CC cardio-vascular diseases (e.g. atherosclerosis), neurodegenerative  
 CC disorders (e.g. Alzheimer's disease and Parkinson's disease), amyotrophic  
 CC lateral sclerosis, apoptosis and inflammatory reactions, allergic  
 CC reactions (e.g. asthma, hay fever and eczema), high bone mass syndrome,  
 CC osteopetrosis, osteoporosis-pseudoglioma syndrome, and Bardet-Biedl  
 CC syndrome 1. The polypeptide is also useful as a specific marker of the  
 CC above diseases or disorders in a wide variety of tissues.  
 CC The discovery of the peroxisome-associated polypeptide  
 CC enables the development of diagnosis and treatment of peroxisomal  
 CC disorders.

XX  
 SQ Sequence 469 BP; 88 A; 132 C; 150 G; 99 T; 0 other;

alignment\_scores:

Quality: 349.00 Length: 90

Ratio: 4.653 Gaps: 1

Percent Similarity: 83.333 Percent Identity: 82.222

alignment\_block:

US-09-486-167A-2 x AAX27967 ..

Align seg 1/1 to: AAX27967 from: 1 to: 469

83 GlyGluTrp.....GlyArgAlaHis 89  
 |||||  
 110 GGAGAGTGGGCTGTGGCGGGTCCGCGAGTTTCACGAGCGCGTGCAGC 159  
 89 sLysAlaGluGlyLysValArgLeuAlaAspProThrGlyAlaPheG 106  
 |||||  
 160 CATGCCCCCAATCAAGGTTCGGCTCTCGCTGATCCACATGGGCGCTTTG 209  
 106 lYlysGluThrAspLeuLeuAspSerLeuValSerIlePheGly 122  
 |||||  
 210 GGAAGGACACAGACTATTACTAGATGATTCGTGTCCATCTTTGGG 259  
 123 AsnArgArgLeuLysArgPheSerMetValValGlnAspGlyIleVally 139  
 |||||  
 260 AATCGACGCTCAAGAGGTTCTCCATGTTGGTACAGGATGGCATAGTAA 309  
 139 sAlaLeuAsnValGluProAspGlyThrGlyLeuThrCysSerLeuAlap 156  
 |||||  
 310 GCCTCTGAATGTGGAACACAGATGGCAGCGCTCACCTGCAGCGCTGGCAC 359  
 156 roAsnIleIleSerGlnLeu 162  
 |||||  
 360 CCATATCATCTCACAGCTC 379

seq\_name: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT:AAx02224

seq\_documentation\_block:

ID\_AAC02224 standard; cDNA; 423 BP.

XX AAC02224;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 2222.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.



PR	01-SEP-1999;	99US-01519130
PR	07-SEP-1999;	99US-01523663
PR	10-SEP-1999;	99US-01530701
PR	13-SEP-1999;	99US-01537078
PR	15-SEP-1999;	99US-01540118
PR	16-SEP-1999;	99US-01540339
PR	20-SEP-1999;	99US-01547759
PR	22-SEP-1999;	99US-01551139
PR	23-SEP-1999;	99US-01551486
PR	24-SEP-1999;	99US-01556559
PR	28-SEP-1999;	99US-01564558
PR	29-SEP-1999;	99US-01565999
PR	04-OCT-1999;	99US-01571117
PR	05-OCT-1999;	99US-01577553
PR	06-OCT-1999;	99US-01578655
PR	07-OCT-1999;	99US-01580209
PR	08-OCT-1999;	99US-01582322
PR	12-OCT-1999;	99US-01583659
PR	13-OCT-1999;	99US-01582933
PR	13-OCT-1999;	99US-01592294
PR	13-OCT-1999;	99US-01592595
PR	14-OCT-1999;	99US-01593229
PR	14-OCT-1999;	99US-01593330
PR	14-OCT-1999;	99US-01593331
PR	14-OCT-1999;	99US-01596377
PR	14-OCT-1999;	99US-01596388
PR	18-OCT-1999;	99US-01595884
PR	21-OCT-1999;	99US-01607411
PR	21-OCT-1999;	99US-01607677
PR	21-OCT-1999;	99US-01607688
PR	21-OCT-1999;	99US-01607770
PR	21-OCT-1999;	99US-01608141
PR	21-OCT-1999;	99US-01608151
PR	22-OCT-1999;	99US-01608150
PR	22-OCT-1999;	99US-01609881
PR	22-OCT-1999;	99US-01609889
PR	25-OCT-1999;	99US-01613404
PR	25-OCT-1999;	99US-01613405
PR	25-OCT-1999;	99US-01614006
PR	26-OCT-1999;	99US-01613359
PR	26-OCT-1999;	99US-01613620
PR	26-OCT-1999;	99US-01613621
PR	26-OCT-1999;	99US-01613631
PR	28-OCT-1999;	99US-01619192
PR	28-OCT-1999;	99US-01619992
PR	28-OCT-1999;	99US-01621391
PR	29-OCT-1999;	99US-01621442
PR	29-OCT-1999;	99US-01621999

```

alignment_block:
US-09-486-167A-2 x AAC33030  ..
Align seg 1/1 to: AAC33030 from: 1 to: 727

1  MetAlaProIleLysValGlyAspAlaIlePro.....AlaValGluVal 15
|||||  |||||||  |||||||  |||||||  |||||||  |||||||
54  ATGGCTCCAATTGTGTGCGCGATGTCGTTCCAGACGGAAGCATTTCTCT 103
|||||  |||||||  |||||||  |||||||  |||||||  |||||||

15  lPhe...GluGlyCyluProGlyAsnLysValAsnLeuAlaGluLeuPheL 31
|||||  |||||||  |||||||  |||||||  |||||||  |||||||

104  CTTGCGATGAGAATGATCAACTTCAGACTGTCTCCGTTCACTCTCTCGCG 153
|||||  |||||||  |||||||  |||||||  |||||||  |||||||

31  ysgLyLysLysGlyValLeuPheGlyValProGlyAlaPheThrProGly 47
|||||  |||||||  |||||||  |||||||  |||||||  |||||||

154  CCGGTAAAAAAGTCATCTCTTTCGGTGTCTCTGGTGTCTTCACTCCCACC 203
|||||  |||||||  |||||||  |||||||  |||||||  |||||||

48  CysSerLysThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLy 64
|||||  |||||||  |||||||  |||||||  |||||||  |||||||

204  TSCAGCATGAAGCATGTTCTCTGGTTTCATTGAGAAAGCAGAGAGCTGAA 253
|||||  |||||||  |||||||  |||||||  |||||||  |||||||

```

64 salalysGlyValGlnValAlaCysLeuSerValAsnAspAlaPheV 81  
 254 GTCAAGGCTGTGATGATCATCTTTGCTTAGTGTGAACGATCCTTTT 303  
 81 alThrGlyGluTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeu 97  
 304 TGATGAAGCATGGGAAGAGACATACCAGAGACAAGCATGTGAAGTTT 353  
 98 LeuAlaAspProThrGlyAlaPhe.....GlyLysGluThrAs 110  
 354 GTAGCTGATGGCTGTGGAGATACACACACACCTTCTTGACTTGAGCTTGA 403  
 110 pleuLeuLeuAspSerLeuValSerIlePheGlyAsnArgArgLeuL 127  
 404 C.....CTTAAGGACAAGGCTGTGGTATT.....AGGTCAA 435  
 127 ysArgPheSerMetValValGlnAspGlyIleValLysAlaLeuAsnVal 143  
 436 GGAGATTCGCTTGTGCTTGAACCTTAAGGTGACTGTAGCCCAATCTC 485  
 144 GluProAspGlyThrGlyLeuThrCysSerLeuAlaProAsnIleIleSe 160  
 486 GAA...TCTGGTGGCGGGTTCACAGATTCCAGCGCGCATGATATCCTCGA 532  
 160 rGlnLeu 162  
 533 GGCTCTC 539

seq\_name: /SIDS1/gcgdata/geneseq/geneseq/NA1995.DAT:AA211145

seq\_documentation\_block:

ID AAT21145 standard; cDNA to mRNA; 315 BP.

AC

XX

XX

XX

DT

XX

XX

DE

XX

XX

XX

OS

XX

XX

PN

XX

PD

XX

PF

XX

XX

XX

PR

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-  
 CC untranslated sequence is unique to a particular mRNA species, almost  
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library  
 CC is constructed so as to reflect accurately the relative abundance of  
 CC different mRNAs in the particular tissue from which it was derived.  
 CC The appearance frequency of a given GS in a cDNA library can be  
 CC determined (esp. using primers and probes derived from the GS  
 CC sequences) as a means of diagnosing abnormal cell function or for  
 CC recognising different cell types.  
 XX  
 SQ Sequence 315 BP; 69 A; 82 C; 81 G; 75 T; 8 other;

alignment\_scores:  
 Quality: 289.00 Length: 62  
 Ratio: 4.898 Gaps: 0  
 Percent Similarity: 95.161 Percent Identity: 93.548

alignment\_block:

US-09-486-167A-2 x AAT21145 ..

Align seg 1/1 to: AAT21145 from: 1 to: 315

100 AspProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSe 116  
 1 GATCCCACTGGGCGCTTGGGAAGNGACAGACTTATNACTAGATGATTC 50  
 116 rLeuValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValV 133  
 51 GCTGGTGTCCATCTTTGGGAATCGACGTCTCAAAAGGTTATCCATGGTGG 100  
 133 alGlnAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGly 149  
 101 TACAGGATGGCATAGTAGAGGCCCTGAATGTGAACACAGATGGCAGGC 150  
 150 LeuThrCysSerLeuAlaProAsnIleLeuSerGln 161  
 151 CTCACCTGCAGCGCTGGCACCAATATCATCTCACAG 186

seq\_name: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT:AAC34962

seq\_documentation\_block:

ID AAC34962 standard; DNA; 737 BP.

XX

XX

XX

XX

DT

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

seq\_name: /SIDS1/gcgdata/geneseq/geneseq/NA1995.DAT:AA211145

seq\_documentation\_block:

ID AAT21145 standard; cDNA to mRNA; 315 BP.

AC

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Identifying gene signatures in 3'-directed human cDNA library - e.g.  
 for diagnosis of abnormal cell function, by preparing cDNA that  
 reflects relative abundance of corresp. mRNA in specific human  
 tissues

Claim 1; Page 813; 2245pp; Japanese.

A single-stranded DNA (or its complementary strand or the corresp.  
 double-stranded DNA) which comprises one of the 7837 "GS" sequences  
 given in AAT19001-T26837 and which is able to hybridise to part of  
 human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)  
 sequences were obtained from 3'-directed cDNA libraries prepared  
 from various human tissues; synthesis of cDNA was initiated from the

Hybridisation assay; genetic mapping; gene expression control;  
 protein identification; signal transduction pathway;  
 metabolic pathway; promoter; termination sequence; ss.

Arabidopsis thaliana.

EPI033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-0301439.

25-FEB-1999; 99US-0121825.

05-MAR-1999; 99US-0123180.

09-MAR-1999; 99US-0123548.

23-MAR-1999; 99US-0125788.

25-MAR-1999; 99US-0126264.

29-MAR-1999; 99US-0126785.

01-APR-1999; 99US-0127462.

06-APR-1999; 99US-0128234.

08-APR-1999; 99US-0128714.

16-APR-1999; 99US-0129845.

19-APR-1999; 99US-0130077.



```
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
  Quality: 289.00      Length: 169
  Ratio: 2.535         Gaps: 6
  Percent Similarity: 67.456  Percent Identity: 40.828

alignment_block:
US-09-486-167a-2 x AAC34962 ..

Align seg 1/1 to: AAC34962 from: 1 to: 737

1 MetAlaProLeuLysValGlyAspAlaLeuPro.....AlaValGluVa 15
62 ATGCTCCAAATGCTGTCGGGATGTCGTACACGCGCACTATCTCCTT 111
15 lPhe...GluGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheL 31
112 CTTCGATGAGAATGATCAACTTTCAGCGCGCTCGTTCACCTCTCGCG 161
31 ysGlyLysLysGlyValLeuPheGlyValProGlyAlaPheThrProGly 47
162 CCGTAAAGAGTCTCTCTTGGTGTCTGCTGCTTTCACCTCCCACT 211
48 CysSerLysThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLy 64
212 TGCAGCATGAAGCATGTTCTCGGTTCATTGAGAACACAGAGGAGCTGAA 261
64 sAlaLysGlyValGlnValValAlaCysLeuSerValAsnAspAlaPheV 81
262 GTCAAGGGTGTGATGAGATCACTTCTTTAGCGTGAACAATCCGTTTG 311
81 alThrGlyGluTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeu 97
312 TGATGAAGCATGGGGAACATFACCAGACAAACATGTGAAGTTT 361
98 LeuAlaAspProThrGlyAlaPhe.....GlyLysGluThrAs 110
362 GTAGCTGATGGTCTGGAAGAATACACACACTTCTCGGACTTCGAGTTGA 411
110 pLeuLeuLeuAspSerLeuValSerIlePheGlyAsnArgLeuL 127
412 C.....CTTAAGACAAAGGTTCTGGTGT.....AGGTCAA 443
127 ysArgPheSerMetValValGlnAspGlyIleValLysAlaLeuAsnVal 143
444 GGAGATTGCTCTGCTACTTGATGATCTCAAGTACTGTAGCCCAATGTC 493
144 GluProAspGlyThrGlyLeuThrCysSerLeuAlaProAsnIleLeu 160
494 GAATCTGGTGGCGAA...TTCACAGTTTCCAGCGCGGATGATCTCTCAA 540
```

```
160 rGlnLeu 162
:
541 GGCTCTC 547

seq_name: /SISL/gcgdata/geneseq/geneseq/NA2000.DAT:AAC45603
seq_documentation_block:
ID AAC45603 standard; DNA; 674 BP.
XX
AC AAC45603;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 47093.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127482.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
```

```
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:      Quality: 288.00      Length: 169
                        Ratio: 2.483          Gaps: 6
Percent Similarity: 68.639      Percent Identity: 40.237
alignment_block:
US-09-486-167A-2 x AAC45603 ..
Align seg 1/1 to: AAC45603 from: 1 to: 674
```

```
*1 MetalaProfilLeysValGlyAspAlailePro.....AlaValGluVa 15
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
54 ATGGCTCCAAATTACTGTGCGGATGTTGTACACGACGAACTACTCTTT 103
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
15 lphe...GluGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheL 31
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
104 CTTGATGAATAATGATCAGCTTCAGACCGTCTCTGTTCACTCTATCGCCG 153
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
31 ysGlyLysLysGlyValLeuPheGlyValProGlyAlaPheThrProGly 47
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
154 CCGGTAAAAAGTCACTCTCTTGGTCTCTCTGCTGCTTCACTCCACA 203
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
48 CysSerLysThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLy 64
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
204 TCCAGCATGACCATGTCCTGATTCATTTGGGAAGCAGAGGAGCTGAA 253
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
64 salalysGlyValGlnValValAlaCysLeuSerValAsnAspAlaPheV 81
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
254 GTCAGAGGATTCATGATGATCTTCTTAGTGTGAATGATCCATTG 303
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
81 alThrGlyGluTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeu 97
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
304 TGATGAAGCATGGGAAAGACATACCAAGAGAACAGCATGTGAAGTTT 353
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
98 LeuAlaaspProThrGlyAlaPhe.....GlyLysGluThrAs 110
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
354 GTAGCAGATGGTGTGGAGAATACACATCTTCTTGACTTGAGCTTGA 403
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
110 pLeuLeuLeuAspSerLeuValSerIlePheGlyAsnArgArgLeuL 127
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
404 T.....CTCAGGACGAGGCTGTGTAT.....AGGTCAA 435
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
127 ysArgPheSerMetValValGlnAspGlyIleValLysAlaLeuAsnVal 143
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
436 GGAGATTGCTCTCTGCTTGATTAACCTTAAGGTAACCTGAGCCAAATGTC 485
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
144 GluProaspGlyThrGlyLeuThrCysSerLeuAlaProAsnIleIleSe 160
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
486 GAA...AATGTTGCGAGTTTCACGGTTTCCAGCGCAGAGGATATCCCTAA 532
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
160 rGlnLeu 162
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
533 GGCTCTC 539
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT.AAC45889
seq_documentation_block:
ID AAC45889 standard; DNA; 489 BP.
XX AAC45889;
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 48136.
XX Arabidopsis thaliana
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EF1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
```

```
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
```

```
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149932.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.

PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
    Quality: 286.00      Length: 169
              Ratio: 2.509      Gaps: 6
Percent Similarity: 67.456      Percent Identity: 40.237

alignment_block:
US-09-486-167A-2 x AAC45889 ..

Align seg 1/1 to: AAC45889 from: 1 to: 489

1 MetAlaProIleLysValGlyAspAlaIlePro.....AlaValGluVal 15
|||||
1 ATGGCTCCAAATTACTGTGCGGATGTTGACACAGCGAAGTATCTCTTT 50
|||||

15 lPhe...GluGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheL 31
|||
51 CTTCGATGAAAAATGATCAGCTTCAGACCGTCTCGGTTCACTCTATCGCG 100
|||||

31 ysGlyLysLysGlyValLeuPheGlyValProGlyAlaPheThrProGly 47
|||||
101 CCGGTAAAAAAGTCATTCTTTGGTGTCTGCTGCTTTCACCTCCACACA 150
|||||

48 CysSerLysThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLy 64
|||||
151 TGCAGTATGAGCCATGTCCTGGATTTCATTCGGGAACGACGAGAGCTGAA 200
|||||

64 sAlaLysGlyValGlnValValAlaCysLeuSerValAsnAspAlaPheV 81
|||||
201 GTCAAGGGTATTGATGAGATCATTTGCTTTAGTGTGAATGATCATTTG 250
|||||

81 alThrGlyGluTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeu 97
|||
251 TGATGAAGGCATGGGAAAACATATCCAGAGAACAAGCATGTGAAGTTT 300
|||||

98 LeuAlaAspProThrGlyAlaPhe.....GlyLysGluThrAs 110
|||||
301 GTACAGATGGTCTGGGAAATACACGCATCTTCTTGACTTGAGCTTGA 350
|||||

110 pLeuLeuAspAspSerLeuValSerIlePheGlyAsnArgArgLeuL 127
|||
351 C.....CTTAAGGACAAAGGGTCTTTGTATT.....AGGTCAA 382
|||||

127 ysArgPheSerMetValValGlnAspGlyIleValLysAlaLeuAsnVal 143
|||||
```

```

343 GGAGATTCCCTTTGTTGCTTGATAACCTTAAGTGACTGTAGCCCAATGTT 432
144 GluProAspGlyThrGlyLeuThrCysSerLeuAlaProAsnIleIleSe 160
   |||      |||      |||      |||      |||      |||      |||      |||
433 GAATCTGGTGGCGAG...TTCACGGTTTCCAGCGCAGAGGATATTCACAA 479
160 rGlnLeu 162
   :      |||
480 GGCTCTT 486

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

```
Run on:      October 22, 2001, 11:20:20 ; Search time 109.78 Seconds
            (without alignments)
            4604.307 Million.cell updated
```

Title: US-09-486-167A-1

Perfect score:

Sequence: 1 gccaggaggcgagtggaag.....ttgtggttgcggaaaaaa 805

Scoring table: IDENTITY NUC

IDENTITY\_NOC  
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

FOSE processing: Minimum Match 0%  
Maximum Match 100%

Maximum Match 100%  
Listing first 45 summaries

```

Database :
N_Geneseq_0601: *
1: /SIDS1/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT.*

```

## SUMMARIES

Result	Query	Score	Match	Length	ID	Description
No.	No.					
1	805	100.0	805	20	AAX27965	Human bronchoalveo
2	798	99.1	993	20	AAX56411	Human vesicle memb
3	462	57.4	604	20	AAX27968	Human bronchoalveo
4	456	56.6	601	20	AAX27966	Human bronchoalveo
5	408.6	50.8	453	21	AAC04049	Human bronchoalveo
6	348.6	43.3	452	22	APF92316	Bovine secreted pro
7	300.8	37.4	469	20	AAX27967	Human mammary tis
8	299	37.1	315	16	AAX21145	Human bronchoalveo
9	209	26.0	423	21	AAC02224	Human gene signatu
10	197	24.5	2710	20	AAX27969	Human secreted pro
11	126.6	15.7	553	21	AAC95054	Cat flea hindgut a

## ALIGNMENTS

RESULT 1

AAX27965  
ID AAX27965 standard; cDNA; 805 BP.

XX  
AC AAX27965:

DT 08-JUN-1999 (first entry)

Human bronchoalveolar polypeptide, B18hum, coding sequence.

B18hum; bronchoalveolar protein; peroxisome-associated polypeptide;  
 lung injury; oxidative stress-related disorder; inflammatory disease;  
 cardiovascular disease; neurodegenerative disorder; allergic reaction;  
 amyotrophic lateral sclerosis; apoptosis; high bone mass syndrome;  
 osteopetrosis; osteoporosis-pseudoglioma syndrome; diagnosis;  
 Bardet-Biedl syndrome 1; therapy; ss.

OS Homo sapiens.

XX

PN- W09909054-A2  
yy

25-FEB-1999.

XX

PF 20-AUG-1998; 98WO-BE00124.  
VY

20-AUG-1997; 97BE-0000692.

PA (UYLO-) UNIV CATHOLIQUE LOU

PA (UYMO-) UNIV MONS-HAINAUT.  
XX

PI Bernard A, Falmagne

XX  
WPT: 1000-180068/T  
DR

XX New low molecular weight human broncho-alveolar polypeptide -  
PT useful for diagnosis and/or treatment of lung injuries and diseases,  
PT and oxidative stress-related diseases and disorders, especially  
PT inflammatory diseases  
XX  
XX Claim 8; Page 33-34; 45pp; English.  
XX  
XX This sequence encodes the human bronchoalveolar polypeptide, designated  
CC B18hum of the invention. B18hum is a low molecular weight human,  
CC peroxisome-associated broncho-alveolar polypeptide. A diagnostic device  
CC featuring the polypeptide, polynucleotide and/or inhibitor is useful for  
CC in vitro detection of lung injuries and diseases or oxidative  
CC stress-related diseases and disorders, especially inflammatory diseases.  
CC The device is also useful for monitoring such diseases or disorders in  
CC patients or fluid samples. The polypeptide, polynucleotide and inhibitor  
CC form pharmaceutical compositions useful in the prevention and/or  
CC treatment of these diseases or disorders, especially specific  
CC cardio-vascular diseases (e.g. atherosclerosis), neurodegenerative  
CC disorders (e.g. Alzheimer's disease and Parkinson's disease), amyotrophic  
CC lateral sclerosis, apoptosis and inflammatory reactions, allergic  
CC reactions (e.g. asthma, hay fever and eczema), high bone mass syndrome,  
CC osteopetrosis, osteoporosis-pseudoglioma syndrome, and Bardet-Biedl  
CC syndrome 1. The polypeptide is also useful as a specific marker of the  
CC above diseases or disorders in a wide variety of tissues.  
CC The discovery of the peroxisome-associated polypeptide  
CC enables the development of diagnosis and treatment of peroxisomal  
CC disorders.  
XX  
XX Sequence 805 BP; 163 A; 200 C; 276 G; 166 T; 0 other;

Query Match 100.0%; Score 805; DB 20; Length 805;  
Best Local Similarity 100.0%; Pred. No. 2.5e-212;  
Matches 805; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 gccaggagcgagtggaagtgccgtggcggtgggactagctggcggtgcgc 60  
Db 1 gccaggagcgagtggaagtgccgtggcggtgggactagctggcggtgcgc 60  
Qy 61 ctgagagcctcagcggtctactcgtcgtgtggcgccggtcagtcagtcgagcg 120  
Db 61 ctgagagcctcagcggtctactcgtcgtgtggcgccggtcagtcagtcgagcg 120  
Qy 121 gcagcaagacggtcagtggaagagagtggtcgtcgtggcggtcccgagttcagcaga 180  
Db 121 gcagcaagacggtcagtggaagagagtggtcgtcgtggcggtcccgagttcagcaga 180  
Qy 181 gccgctgcagccatggcccaatcaagtggtggagatgccatcccgagtcagtggttt 240  
Db 181 gccgctgcagccatggcccaatcaagtggtggagatgccatcccgagtcagtggttt 240  
Qy 241 gaaggagcagcaggaagagagtggtgagcagagtcgttcaaggcagaaggggtg 300  
Db 241 gaaggagcagcaggaagagagtggtgagcagagtcgttcaaggcagaaggggtg 300  
Qy 301 ctgtttgagttcctggggtccctcagcgtggttcccaagacacacctgccaggggtt 360  
Db 301 ctgtttgagttcctggggtccctcagcgtggttcccaagacacacctgccaggggtt 360  
Qy 361 gtgagcagcgtgagtcgtgagcagcagcagcagtcagtcgtggtcgtcagtggt 420  
Db 361 gtgagcagcgtgagtcgtgagcagcagcagcagtcagtcgtggtcgtcagtggt 420  
Qy 421 aatgatccttctgactggtcagcagtggtggcgagcccaagcagcagcagcaggttcgg 480  
Db 421 aatgatccttctgactggtcagcagtggtggcgagcccaagcagcagcagcaggttcgg 480  
Qy 481 ctctggtgatcccatcctggggtccttgggaaggagacagacttattactagatgattcg 540  
Db 481 ctctggtgatcccatcctggggtccttgggaaggagacagacttattactagatgattcg 540  
Qy 541 ctgggtgtccatcttgggaatcagctctcaagaggttctccatcgtggttacagagtggc 600

Db 541 ctgggtgtccatcttgggaatcagctctcaagaggttctccatcgtggtgagagtggc 600  
Qy 601 atagtgaagccctggaatgtggaaccagatgacagcagcctcagctgagcagcagc 660  
Db 601 atagtgaagccctggaatgtggaaccagatgacagcagcctcagctgagcagcagc 660  
Qy 661 aatatcatctcacagctcgtgagccctggccagatctacttctccaccctccctatct 720  
Db 661 aatatcatctcacagctcgtgagccctggccagatctacttctccaccctccctatct 720  
Qy 721 cactgtccagccctgtgctggccctgcaatggaaatgttggccagatttctgcaata 780  
Db 721 cactgtccagccctgtgctggccctgcaatggaaatgttggccagatttctgcaata 780  
Qy 781 aacactgtggttgcgaaaaaaa 805  
Db 781 aacactgtggttgcgaaaaaaa 805  
RESULT 2  
AA56411  
ID AAX56411 standard; cDNA; 993 BP.  
XX AC  
XX AAX56411;  
DT 22-JUL-1999 (first entry)  
XX  
XX Human vesicle membrane protein-like protein 1 encoding cDNA.  
XX  
XX Human; vesicle membrane protein-like protein; VMP; Cushing's syndrome;  
KW developmental disorder; vesicle-trafficking disorder; cystic fibrosis;  
KW immunological disorder; reproductive disorder; neoplastic disorder;  
KW anaemia; muscular dystrophy; cataract; Grave's disease; allergy;  
KW ulcerative colitis; microbial infection; Addison's disease; cancer; ss.  
XX  
XX Homo sapiens.  
XX OS  
XX WO9921994-A2.  
XX  
XX 06-MAY-1999.  
XX  
XX 14-OCT-1998; 98WO-US21730.  
XX  
XX 28-OCT-1997; 97US-0959004.  
XX  
XX (INCY-) INCYTE PHARM INC.  
XX  
XX Corley NC, Hillman JL, Lal P, Shah P, Yue H;  
XX WPI; 1999-326702/27.  
XX P-PSDB; AAY17388.  
XX  
XX Human vesicle membrane-like proteins, useful for diagnosis,  
XX treatment and prevention of e.g. developmental disorders  
XX  
XX Claim 7; Fig 1; 105pp; English.  
XX  
XX The present sequence encodes human vesicle membrane protein-like  
XX protein 1 (VMP-1). VMP proteins, and their fragments, are used to treat  
XX or prevent developmental or vesicle-trafficking disorders, while their  
XX antagonists are used to treat or prevent immunological, reproductive or  
XX neoplastic disorders. Typical of many such disorders are anaemia;  
XX Cushing's syndrome; muscular dystrophy; cataract; cystic fibrosis;  
XX Grave's disease; ulcerative colitis; allergies; microbial infections;  
XX Addison's disease; cancer of breast, testis and prostate. VMP proteins  
XX are also used to raise specific antibodies (used to detect VMP in  
XX immunoassays (for diagnosis or monitoring), in competitive drug screens  
XX and to purify VMP from natural sources) and to screen for specific  
XX antagonists (potential therapeutic agents). VMP polynucleotides or their  
XX fragments, are used in hybridization assays to detect VMP in biological  
XX samples (e.g. for diagnosis, including detection of mutations and  
XX polymorphisms), optionally after amplification; to express recombinant





Db 346 ggaacagacttattactagatgattcgctgggtgtccatcttttgggaatcgacgtctcaa 405  
 QY 573 gagttctccatggtgtacagatgcatagtgaaagccctgaatgtgaaacagatgg 632  
 Db 406 gagttctccatggtgtacagatgcatagtgaaagccctgaatgtgaaacagatgg 465  
 QY 633 cacaggcctcaactgacgtgacccccaataatcattcaccagctctgagggccctggcc 692  
 Db 466 cacaggcctcaactgacgtgacccccaataatcattcaccagctctgagggccctggcc 525  
 QY 693 agattacttctccacccctccctattctccctgcccagccctgtgtggtggccctgcaa 752  
 Db 526 agattacttctccacccctccctattctccctgcccagccctgtgtggtggccctgcaa 585  
 QY 753 ttggaattgtggccag 768  
 Db 586 ttggaattgtggccag 601

RESULT 5  
 AAC04049  
 ID AAC04049 standard; cDNA; 453 BP.  
 AC AAC04049;  
 XX  
 DT 06-OCT-2000 (first entry)  
 DE Human secreted protein 5' EST, SEQ ID NO: 4047.  
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping; ss.  
 XX Homo sapiens.  
 OS  
 PN EP1033401-A2.  
 PD 06-SEP-2000.  
 XX  
 PF 21-FEB-2000; 2000EP-0200610.  
 XX  
 PR 26-FEB-1999; 99US-0122487.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Dumas Milne Edwards J, Duclert A, Giordano J;  
 XX  
 DR WPI; 2000-500381/45.  
 DR P-PSDB; AAG04043.  
 XX  
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 XX  
 PS Claim 1; SEQ ID 4047; 71pp + CD-ROM; English.  
 CC  
 CC The present sequence is one of a large number of 5' ESTs derived from  
 CC mRNAs encoding secreted proteins. An ORF has been identified within the  
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs  
 CC derived from 30 different tissues. EST sequences usually correspond  
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are  
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not  
 CC well suited for isolating cDNA sequences derived from the 5' ends of  
 CC mRNAs and even in those cases where longer cDNA sequences have been  
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from  
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length  
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
 CC gene therapy and chromosome mapping procedures. They are used to obtain  
 CC upstream regulatory sequences and to design expression and secretion  
 CC vectors.  
 XX  
 SQ Sequence 453 BP; 82 A; 108 C; 180 G; 82 T; 1 other;

Query Match 50.8%; Score 408.6; DB 21; Length 453;  
 Best Local Similarity 99.5%; Pred. No. 3.2e-103;  
 Matches 419; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
 QY 1 gccagagagcgcagtggaagtgcctgtggcgccgtatggagactagctgagctgtgcgcc 60  
 Db 34 gccagagagcgcagtggaagtgcctgtggcgccgtatggagactagctgagctgtgcgcc 93  
 QY 61 ctgagacgctcagcgggctatatactcgtggtggcgccggtcagctctgagcgagcg 120  
 Db 94 ctgagacgctcagcgggctatatactcgtggtggcgccggtcagctctgagcgagcg 153  
 QY 121 gcagcaagacgctgcagtgaaagagagtggtggtggtggtggtggtggtggtggtggt 180  
 Db 154 gcagcaagacgctgcagtgaaagagagtggtggtggtggtggtggtggtggtggtggt 213  
 QY 181 gccgctgcagccatgcccccaatcaagtggtggtggtggtggtggtggtggtggtggt 240  
 Db 214 gccgctgcagccatgcccccaatcaagtggtggtggtggtggtggtggtggtggtggt 273  
 QY 241 gaaggaggagccaggaacagtgtaacctggcagagctgttcaagggaagaggtgtgtg 300  
 Db 274 gaaggaggagccaggaacagtgtaacctggcagagctgttcaagggaagaggtgtgtg 333  
 QY 301 ctgtttgaggttctcgtgggcttccacctggtatgttccaaagacacacacgtgccagggttt 360  
 Db 334 ctgtttgaggttctcgtgggcttccacctggtatgttccaaagacacacacgtgccagggttt 393  
 QY 361 gtcgagcagcgtgcagctctgaaggccaaggtggtggtggtggtggtggtggtggtggt 420  
 Db 394 gtcgagcagcgtgcagctctgaaggccaaggtggtggtggtggtggtggtggtggtggt 452  
 QY 421 a 421  
 Db 453 a 453

RESULT 6  
 AAF92316  
 ID AAF92316 standard; cDNA; 452 BP.  
 XX  
 AC AAF92316;  
 XX  
 DT 15-MAY-2001 (first entry)  
 DE Bovine mammary tissue derived cDNA #29.  
 KW Bovine; mammary gland; cancer; tumour; angiogenesis; ss.  
 OS Bos taurus.  
 XX  
 PN WO200114553-A1.  
 XX  
 PD 01-MAR-2001.  
 XX  
 PF 23-AUG-2000; 2000WO-NZ00166.  
 XX  
 PR 23-AUG-1999; 99US-0150330.  
 XX  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (NZPA-) NEW ZEALAND PASTORAL AGRIC RES INST LTD.  
 XX  
 PI Havukkala IJ, Gleen M, Grigor MR, Molenaar AJ;  
 XX  
 DR WPI; 2001-226619/23.  
 XX  
 PT New polypeptides and polynucleotides encoding the polypeptides, which  
 PT are expressed in bovine mammary gland tissue, useful for stimulating  
 PT mammary gland growth or function, or inducing differentiation of milk  
 PT producing cells -  
 XX  
 PS Claim 1; Page 49-50; 97pp; English.  
 XX





Qy 1 gccaggagcagtggaagtggccgtggcggtatgggaactagctggcgtgccc 60  
 Db 34 gccaggagcagtggaagtggccgtggcggtatgggaactagctggcgtgccc 93  
 Qy 61 ctgagacgtcagcggtctatctgctggtggcggtcgcgtcagctcgcgcagcg 120  
 Db 94 ctgagacgtcagcggtctatctgctggtggcggtcgcgtcagctcgcgcagcg 153  
 Qy 121 gcagcaagacggtgcagtggaaggagtggtggcgtctggtggcggtccgcagtttcagcaga 180  
 Db 154 gcagcaagacggtgcagtggaaggagtggtggcgtctggtggcggtccgcagtttcagcaga 213  
 Qy 161 gccgtgcagcagtggtcccaatcaagt 209  
 Db 214 gccgtgcagcagtggtcccaatcaagt 242

RESULT 10  
 AAX27969  
 ID AAX27969 standard; cDNA; 2710 BP.  
 XX AC AAX27969;  
 XX DT 08-JUN-1999 (first entry)  
 XX DE Human bronchoalveolar polypeptide, B18hum, coding sequence.  
 XX KW B18hum; bronchoalveolar protein; peroxisome-associated polypeptide;  
 KW lung injury; oxidative stress-related disorder; inflammatory disease;  
 KW cardiovascular disease; neurodegenerative disorder; allergic reaction;  
 KW amyotrophic lateral sclerosis; apoptosis; high bone mass syndrome;  
 KW osteopetrosis; osteoporosis-pseudoglioma syndrome; diagnosis;  
 KW Bardet-Biedl syndrome 1; therapy; ss.  
 XX OS Homo sapiens.  
 XX PN WO9909054-A2.  
 XX PD 25-FEB-1999.  
 XX PF 20-AUG-1998; 98WO-BE00124.  
 XX PR 20-AUG-1997; 97BE-0000692.  
 XX PA (UYLO-) UNIV CATHOLIQUE LOUVAIN.  
 XX PA (UYMO-) UNIV MONS-HAINAUT.  
 XX PI Bernard A, Falmagne P, Hermans C, Knoop B, Wattiez R;  
 XX WI 1999-180968/15.  
 XX PT New low molecular weight human broncho-alveolar polypeptide -  
 PT useful for diagnosis and/or treatment of lung injuries and diseases,  
 PT and oxidative stress-related diseases and disorders, especially  
 PT inflammatory diseases  
 XX Example 6; Page 42-43; 45pp; English.  
 XX This sequence encodes the human bronchoalveolar polypeptide, designated  
 CC B18hum of the invention. B18hum is a low molecular weight human,  
 CC peroxisome-associated broncho-alveolar polypeptide. A diagnostic device  
 CC featuring the polypeptide, polynucleotide and/or inhibitor is useful for  
 CC in vitro detection of lung injuries and diseases or oxidative  
 CC stress-related diseases and disorders, especially inflammatory diseases.  
 CC The device is also useful for monitoring such diseases or disorders in  
 CC patients or fluid samples. The polypeptide, polynucleotide and inhibitor  
 CC form pharmaceutical compositions useful in the prevention and/or  
 CC treatment of these diseases (e.g. atherosclerosis), neurodegenerative  
 CC cardio-vascular diseases (e.g. Alzheimer's disease and Parkinson's disease), amyotrophic  
 CC lateral sclerosis, apoptosis and inflammatory reactions, allergic  
 CC reactions (e.g. asthma, hay fever and eczema), high bone mass syndrome,

CC osteopetrosis, osteoporosis-pseudoglioma syndrome, and Bardet-Biedl  
 CC syndrome 1. The polypeptide is also useful as a specific marker of the  
 CC above diseases or disorders in a wide variety of tissues.  
 CC The discovery of the peroxisome-associated polypeptide  
 CC enables the development of diagnosis and treatment of peroxisomal  
 CC disorders.  
 XX SQ Sequence 2710 BP; 550 A; 748 C; 767 G; 645 T; 0 other;  
 Query Match 24.5%; Score 197; DB 20; Length 2710;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-44;  
 Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 574 aggttctccatggtgtacagatggtacagatggaaggccctgaatgtggaaccagatggc 633  
 Db 2514 aggttctccatggtgtacagatggtacagatggaaggccctgaatgtggaaccagatggc 2573  
 Qy 634 acaggcctcactgcagcctggcaccacaaatcatctcacagctctgagggccctggcca 693  
 Db 2574 acaggcctcactgcagcctggcaccacaaatcatctcacagctctgagggccctggcca 2633  
 Qy 694 gattactctccaccctccctctatctcacctgcccagccctgtgtggtggccctgcaat 753  
 Db 2634 gattactctccaccctccctctatctcacctgcccagccctgtgtggtggccctgcaat 2693  
 Qy 754 tggaatgtggccagat 770  
 Db 2694 tggaatgtggccagat 2710

RESULT 11  
 AAC95054  
 ID AAC95054 standard; cDNA; 553 BP.  
 XX AC AAC95054;  
 XX DT 19-FEB-2001 (first entry)  
 XX DE Cat flea hindgut and Malpighian tubule (HMT) cDNA, SEQ ID NO:1549.  
 XX KW Cat flea; hindgut and Malpighian tubule nucleic acid; HMT;  
 KW flea infestation; vaccine; antiparasitic; therapeutic target;  
 KW diagnosis; detection; ss.  
 XX OS Ctenocephalides felis.  
 XX PN WO2000061621-A2.  
 XX PD 19-OCT-2000.  
 XX PF 07-APR-2000; 2000WO-US09437.  
 XX PR 09-APR-1999; 99US-0128704.  
 XX PA (HESK-) HESKA CORP.  
 XX PI Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N;  
 XX WI 2000-656323/63.  
 XX PT Flea Malpighian tubule and head and nerve cord tissue derived nucleic  
 PT acids useful for the prevention, diagnosis and treatment of flea  
 PT infestations -  
 XX Claim 26; Page 756; 964pp; English.  
 XX The invention relates to novel cat flea (Ctenocephalides felis) nucleic  
 CC acids which are expressed in hindgut and Malpighian tubule (HMT) tissue  
 CC or head and nerve cord (HNC) tissue. The invention also relates to the  
 CC encoded proteins. The invention additionally encompasses expression  
 CC constructs, recombinant viruses and recombinant cells comprising the  
 CC nucleic acids of the invention, recombinant production of the proteins,





XX Arabidopsis thaliana DNA fragment SEQ ID NO: 8494.  
DE XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 08-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134236.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.





Db 447 acaaacatgtgaagttgttagctgtagtggtctgag 482

## RESULT 15

AAC34962

ID AAC34962 standard; DNA; 737 BP.

XX

AC AAC34962;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 8523.

XX

KW Hybridisation assay; genetic mapping; gene expression control;  
protein identification; signal transduction pathway;  
metabolic pathway; promoter; termination sequence; ss.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR

PR 05-MAR-1999; 99US-0123180.

PR

PR 09-MAR-1999; 99US-0123548.

PR

PR 23-MAR-1999; 99US-0125788.

PR

PR 25-MAR-1999; 99US-0126284.

PR

PR 29-MAR-1999; 99US-0126785.

PR

PR 01-APR-1999; 99US-0127462.

PR

PR 06-APR-1999; 99US-0128234.

PR

PR 08-APR-1999; 99US-0128714.

PR

PR 16-APR-1999; 99US-0129845.

PR

PR 19-APR-1999; 99US-0130077.

PR

PR 21-APR-1999; 99US-0130449.

PR

PR 23-APR-1999; 99US-0130510.

PR

PR 28-APR-1999; 99US-0130891.

PR

PR 30-APR-1999; 99US-0131449.

PR

PR 30-APR-1999; 99US-0132048.

PR

PR 04-MAY-1999; 99US-0132484.

PR

PR 05-MAY-1999; 99US-0132485.

PR

PR 06-MAY-1999; 99US-0132486.

PR

PR 07-MAY-1999; 99US-0132487.

PR

PR 11-MAY-1999; 99US-0132863.

PR

PR 14-MAY-1999; 99US-0132566.

PR

PR 14-MAY-1999; 99US-0134218.

PR

PR 14-MAY-1999; 99US-0134219.

PR

PR 14-MAY-1999; 99US-0134221.

PR

PR 18-MAY-1999; 99US-0134370.

PR

PR 18-MAY-1999; 99US-0134768.

PR

PR 19-MAY-1999; 99US-0134941.

PR

PR 20-MAY-1999; 99US-0135124.

PR

PR 21-MAY-1999; 99US-0135353.

PR

PR 24-MAY-1999; 99US-0135622.

PR

PR 25-MAY-1999; 99US-0136021.

PR

PR 27-MAY-1999; 99US-0136392.

PR

PR 28-MAY-1999; 99US-0136782.

PR

PR 01-JUN-1999; 99US-0137222.

PR

PR 03-JUN-1999; 99US-0137528.

PR

PR 04-JUN-1999; 99US-0137502.

PR

PR 07-JUN-1999; 99US-0137724.

PR

PR 08-JUN-1999; 99US-0138094.

PR

PR 10-JUN-1999; 99US-0138540.

PR

PR 10-JUN-1999; 99US-0138847.

PR

PR 14-JUN-1999; 99US-0139119.

PR

PR 16-JUN-1999; 99US-0139452.

PR

PR 16-JUN-1999; 99US-0139453.

PR

PR 17-JUN-1999; 99US-0139492.

PR

PR 18-JUN-1999; 99US-0139454.

PR

PR 18-JUN-1999; 99US-0139455.

PR

PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.





OM of: 'US-09-486-167A-2 to: GenEmbl:\* out\_format : pfs

Date: Oct 22, 2001 10:41 AM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=frame; p2n.model -DEV=xlp  
-O/cgn2.1/USPTO.spool/US09486167/runat\_22102001\_064207\_25024/app\_query.fasta\_1.222  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500  
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0  
-MAXLEN=2000000000 -USER=US09486167 @CGL\_1.4855 -NCPU=6  
-ICPU=3 -LONGLOG -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-486-167A-2

Query length: 182

Database: GenEmbl.\*

Database sequences: 1344157

Database length: -856060004

Search time (sec): 1278.290000

score\_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_pr9:HSA249483	+	830.00	1418.11	9.4e-71	489	!
gb_pr5:AF231705	+	830.00	1414.28	1.5e-70	788	!
gb_pat1:A99097	+	830.00	1414.11	1.6e-70	805	!
gb_pr4:AF110731	+	830.00	1414.11	1.6e-70	805	!
gb_pr5:AF112212	+	830.00	1413.82	1.6e-70	835	!
gb_pr5:AF124993	+	830.00	1413.77	1.6e-70	840	!
gb_pr5:AF242525	+	830.00	1413.44	1.7e-70	875	!
gb_pr4:AF110736	+	822.00	1401.25	8.2e-70	725	!
gb_pr5:AF197952	+	820.00	1397.24	1.4e-69	780	!
gb_pr4:AF110734	+	818.00	1394.40	1.3e-69	725	!
gb_pat1:A99101	+	767.00	1307.66	1.3e-64	675	!
gb_r01:AF110733	+	767.00	1307.66	1.3e-64	675	!
gb_r01:AF197951	+	767.00	1307.51	1.4e-64	688	!
gb_r01:AF110732	+	767.00	1305.93	1.7e-64	837	!
gb_r01:AF124994	+	767.00	1305.44	1.8e-64	895	!
gb_r01:AF110735	+	761.00	1296.04	5.9e-64	798	!
gb_r01:AF110735	+	752.00	1280.82	4.2e-63	881	!
gb_r01:AF305564	+	747.00	1271.28	1.4e-62	881	!
gb_r01:AF305564	+	597.50	1009.29	5.6e-48	1869	!
gb_r01:AF305564	+	573.00	976.44	3.8e-46	601	!
gb_r01:AF305564	+	568.00	967.84	1.1e-45	604	!
gb_r01:AF305564	+	498.50	837.92	2.0e-38	2360	!
gb_r01:AF305564	+	486.00	783.13	2.2e-35	153747	!
gb_r01:AF305564	+	486.00	782.29	2.5e-35	168475	!
gb_r01:AF305564	+	486.00	782.00	2.5e-35	174562	!
gb_r01:AF305564	+	476.50	799.14	2.8e-36	2710	!
gb_r01:AF305564	+	463.50	789.40	9.8e-36	569	!
gb_r01:AF305564	+	424.50	690.25	3.3e-30	157974	!
gb_r01:AF305564	+	424.50	677.84	1.6e-29	32270	!
gb_r01:AF305564	+	417.00	662.97	1.1e-28	151611	!
gb_r01:AF305564	+	406.00	647.57	7.8e-28	224896	!
gb_r01:AF305564	+	395.00	633.97	4.5e-27	195316	!
gb_r01:AF305564	+	378.50	639.52	6.2e-27	127234	!
gb_r01:AF305564	+	373.50	631.00	2.2e-27	66266	!
gb_r01:AF305564	+	349.00	594.91	6.7e-25	979	!
gb_r01:AF305564	+	345.00	544.81	4.1e-22	1438886	!
gb_r01:AF305564	+	321.50	545.85	3.6e-22	1439722	!
gb_r01:AF305564	+	321.50	544.39	4.4e-22	979	!

gb\_om:S65367 - 314.00 523.14 6.7e-21 2056 ! S65367 leucine aminopeptidase  
gb\_ba1:AE006000 - 312.50 506.22 5.8e-20 12299 ! AE006000 Caulobacter crescentii  
gb\_ba2:AF003012 - 310.50 476.03 2.8e-18 346547 ! AF003012 Mesorhizobium lotum  
gb\_pl1:AF133302 + 300.00 507.60 4.9e-20 718 ! AF133302 Brassica rapa subsp. chinensis  
seq\_name: gb\_pr9:HSA249483  
seq\_documentation\_block:  
LOCUS HSA249483 489 bp mRNA PRI 02-DEC-1999  
DEFINITION Homo sapiens mRNA for human thiol peroxidase homologous protein (TPX VI gene).  
ACCESSION AJ249483  
VERSION AJ249483.1 GI:6523288  
KEYWORDS human thiol peroxidase homologous protein; TPX VI gene.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Kim, I. H. and Jeong, W.  
TITLE A new type of human thiol peroxidase (Human TPX Type VI)  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 489)  
AUTHORS Kim, I. H.  
TITLE Direct Submission  
JOURNAL Submitted (13-SEP-1999) Kim I.H., Biochemistry, Paichai University, 439-6 Doma-2-Dong Seo-Gu, Taejeon 302-735, REPUBLIC OF KOREA  
FEATURES  
Location/Qualifiers  
1..489  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_line="HeLa Cell"  
1..489  
/gene="TPX VI"  
1..489  
/gene="TPX VI"  
/function="peroxidase"  
/codon\_start=1  
/evidence="experimental"  
/product="human thiol peroxidase homologous protein"  
/protein\_id="CA682210.1"  
/db\_xref="GI:6523288"  
/translation="MAPKVGDAIPAVEVEFEGEPGNKVNLAELFKGKGVLFQVPGAF  
TPGSKTHLPGEVQAEALKAQVQVACLVSNDFAVTVGEMRAHKAQKQVLLADPT  
GAFKETDLDDSLVSIFGNRLKRFNSVWDGIVKALNVEPDGTGLTCLAPNLIS  
QL"  
BASE COUNT 107 a 114 c 163 g 105 t  
ORIGIN  
alignment\_scores:  
Quality: 830.00 Length: 162  
Ratio: 5.123 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-486-167A-2 x HSA249483  
Align seg 1/1 to: HSA249483 from: 1 to: 489  
1 MetAlaProTLeLysValGlyAspAlaAlaLeuProAlaValGluValPheG1 17  
|||||  
1 ATGCCCCCAATCAAGTGGGAGATGTCATCCAGCAGCAGTGGAGTGTGTTGA 50  
17 uGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34  
|||||  
51 AGGGGAGCCAGGGAACAAGTGAACCTGGCAGAGCTGTTCAAGGCGAAGA 100  
34 yscLlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
|||||  
101 AGGGT 150  
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 57

Mon Oct 22 14:46:56 2001

us-09-486-167a-2.rge

151 ACACACCTGCCAGGCTTTGGAGCAGGCTGAGGCTCTGAAGGCCAAGG 200  
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84  
201 AGTCCAGGTGGTGGCTGCTGAGTGTAAATGATGCTTTGTGACTGGCG 250  
84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100  
251 AGTGGGCCGAGCCCAACAGCGGAGGCAAGGTTCCGGCTCTCGCTGAT 300  
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117  
301 CCCACTGGGCGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCGCT 350  
117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134  
351 GGTGTCCTCTTTGGGAATCGAGCTCTCAAGAGGTTCTCATGGTGTAC 400  
134 lnAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150  
401 AGGATGGCCTAGTGAAGGCCCTCAATGTGAACACAGATGGCACAGGCTC 450  
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162  
451 ACCTGCAGCTGGCAGCCCAATATCATCTCACAGCTC 486

seq\_name: gb\_pr5:AF231705

seq\_documentation\_block: 788 bp mRNA PRI 27-JUN-2000  
LOCUS AF231705 Homo sapiens Alu co-repressor 1 (ACR1) mRNA, complete cds.  
ACCESSION AF231705  
VERSION AF231705.1 GI:8745393

KEYWORDS  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 788)  
AUTHORS Kropotov,A., Sedova,V., Ivanov,V., Sazeeva,N., Tomilin,A.,  
Krutillina,R., Oei,S.L., Griesenbeck,J., Buchlow,G. and Tomilin,N.  
A novel human DNA-binding protein with sequence similarity to a  
subfamily of redox proteins which is able to repress  
RNA-polymerase-III-driven transcription of the Alu-family  
retroposons in vitro  
Eur. J. Biochem. 260 (2), 336-346 (1999)

JOURNAL MEDLINE 99195471  
PUBMED 10095767  
REFERENCE 2 (bases 1 to 788)  
AUTHORS Kropotov,A.V. and Tomilin,N.V.  
TITLE Direct Submission  
JOURNAL Submitted (04-FEB-2000) Chromosome Stability, Institute of Cytology  
RAS, Tikhoretskii Av.4, St.Petersburg 194064, Russia

FEATURES  
source Location/Qualifiers  
1..788  
/organism="Homo sapiens"  
/db\_xref="taxon.9606"  
/chromosome="11"  
/map="11q13"  
1..788  
/gene="ACR1"  
20..664  
/gene="ACR1"  
/codon\_start=1  
/product="Alu co-repressor 1"  
/protein\_id="AAF78899.1"  
/db\_xref="GI:8745394"  
/translation="MGLAVCALRRSAGYILVGGAGGQSAARRCSEGEWAGGVGR  
SFSPAAAMAPIKVGDALPAVEVEPEPKNVLAELFKGKGVLFVGPVGAFTPGCSK  
THLPFGVEQAEAKGVQVACISVNDFAVTGEWGRAHKAEGKRVRLADPTGAFGRE  
TDLLEDSLSIFSNRRRLKRFSEVMVQDGIKVALNVEPDGTGLTCSLAPNIISQL"

gene  
CDS

BASE COUNT 153 a 201 c 268 g 166 t

ORIGIN

alignment\_scores: Quality: 830.00 Length: 162  
Ratio: 5.123 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-486-167A-2 x AF231705 ..  
Align seg 1/1 to: AF231705 from: 1 to: 788  
1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG 17  
176 ATGGCCCCAATCAAGTGGGAGATGCATCCACGACAGTGGAGGTGTTGA 225  
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34  
226 AGGGAGCCAGGGAACCTGACCTGGCAGAGCTGTTCAAGGGCAAGA 275  
34 yGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
276 AGGGTGTGCTGTTGGAGTTCCTGGGGCCTTCACCCCTGGATGTTCCAAG 325  
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG 67  
326 ACACACCTGCCAGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGG 375  
67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84  
376 AGTCCAGGTGGTGGCCTGCTGTAGTGTAAATGATGCTTTGTGACTGGCG 425  
84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100  
426 AGTGGGCCGAGCCCAACAGCGGAGGCAAGGTTTCGGCTCTCGCTGAT 475  
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117  
476 CCCACTGGGCGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCCT 525  
117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134  
526 GGTGTCCATCTTTGGGAATCCAGCTCTCAAGAGGTTCTCCATGGTGTAC 575  
134 lnAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150  
576 AGGATGGCCTAGTGAAGGCCCTCAATGTGAACACAGATGGCACAGGCTC 625  
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162  
626 ACCTGCAGCTGGCAGCCCAATATCATCTCACAGCTC 661

seq\_name: gb\_pat1:A99097

seq\_documentation\_block: 805 bp DNA PAT 20-SEP-2000  
LOCUS A99097  
DEFINITION Sequence 1 from Patent WO9909054.  
ACCESSION A99097  
VERSION A99097.1 GI:6782047  
KEYWORDS  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 805)  
AUTHORS Falmagne,P., Wattiez,R., Bernard,A., Hermans,C. and Knoops,B.  
TITLE peroxisome-associated polypeptide, nucleotide sequence encoding  
said polypeptide and their uses in the diagnosis and/or the  
treatment of lung injuries and diseases, and of oxidative  
stress-related disorders

JOURNAL Patent: WO 9909054-A 1 25-FEB-1999;  
UNIV MONS HAINAUT (BE) ; FALMAGNE PAUL (BE) ; WATTIEZ RUDDY (BE) ;

BERNARD ALFRED (BE) ; HERMANS CEDRIC (BE) ; KNOOPS BERNARD (BE) ;  
UNIV LOUVAIN (BE)  
Location/Qualifiers  
1. .805  
/organism="Homo sapiens"  
193. .681  
/note="unnamed protein product"  
/codon\_start=1  
/db\_xref="GI:6782048"  
/translation="MAPIRKVGDAIPAVEFEGEPGNKNVLAELFRGKGVLFVGPAGF  
TPGSKTHLPFGVEQAEALKAGVQVACLSDAFVTCGEMRAHKAEGKVRLLADPT  
GAPGKETDLDLSDLSVIFGNRLKRFMSVVDGIVKALNVEPDGTGLTCSLAPNIS  
QL"  
BASE COUNT 163 a 200 c 276 g 166 t  
ORIGIN

alignment\_scores:  
Quality: 830.00 Length: 162  
Ratio: 5.123 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-486-167A-2 x A99097 ..  
Align seg 1/1 to: A99097 from: 1 to: 805

1 MetAlaProLysValGlyAspAlaIleProAlaValGluValPheG1 17  
193 ATGCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTGA 242

17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysLysL 34  
243 AGGGACCCAGGGAACAAAGTGAACCTGGCAGAGCTGTTCAAGGCAAGA 292

34 ySGlyValLeuPheGlyValProGlyValAlaPheThrProGlyCysSerLys 50  
293 AGGTGTGCTGTTGGAGTTCCTGGGGCTTCACCCCTGGATGTTCCAAG 342

51 ThrHisLeuProGlyPheValGluGlnAlaLeuLysAlaLysG1 67  
343 ACACACTGCCAGGTTTGTGGAGCAGCTGAGGCTCTCAAGGCCAAGG 392

67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84  
393 AGTCCAGGTGGTCTGCTGAGTGTATGATGCTTGTGACTGGCG 442

84 LuTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100  
443 AGTGGGGCCGACCCACAAAGCGGAAGCAAGTTCGGCTCTCGCTGAT 492

101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117  
493 CCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCGCT 542

117 uValSerLeuPheGlyAsnArgArgLeuLysArgPheSerMetValValG 134  
543 GGTGTCCATCTTTGGGAATGACGCTCTCAAGAGGTTCCTCATGGTGGTAC 592

134 InAspGlyLeuValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150  
593 AGGATGGCATAGTGAAGGCCCTGAATGTGGAACCAAGATGGCACAGGCCTC 642

151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162  
643 ACCTGCAGCCCTGGCACCACCAATATCATCTCACAGCTC 678

seq\_name: gb\_pr4:AF110731  
seq\_documentation\_block:  
LOCUS AF110731 805 bp mRNA PRI 26-SEP-2000

DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Homo sapiens antioxidant enzyme B166 mRNA, complete cds.  
AF110731  
AF110731.1 GI:6103723  
human.  
Homo sapiens

REFERENCE  
AUTHORS  
TITLE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 805)  
Wattiez,R., Hermans,C., Bernard,A., Lesur,O. and Falmagne,P.  
Human bronchoalveolar lavage fluid: two-dimensional gel  
electrophoresis, amino acid microsequencing and identification of  
major proteins  
Electrophoresis 20 (7), 1634-1645 (1999)

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

99351776  
1042490  
2 (bases 1 to 805)  
Knoops,B., Clippe,A., Bogard,C., Arsalane,K., Wattiez,R.,  
Hermans,C., Duconseille,E., Falmagne,P. and Bernard,A.  
Cloning and characterization of AOEB166, a novel mammalian  
antioxidant enzyme of the peroxiredoxin family  
J. Biol. Chem. 274 (43), 30451-30458 (1999)  
99452929  
10521424  
3 (bases 1 to 805)  
Knoops,B., Wattiez,R., Falmagne,P., Hermans,C. and Bernard,A.  
Direct Submission  
Submitted (04-DEC-1998) Department of Biology, Catholic University  
of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium  
Location/Qualifiers  
1. .805  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11q13; about 7cR from marker D11S913; between  
markers D11S1963 and D11S4407"  
/tissue\_type="lung"  
37. .681  
/note="hAOEB166; alkyl hydroperoxide reductase;  
thioredoxin peroxidase; represents prxv peroxiredoxin  
subfamily in mammals"  
/codon\_start=1  
/product="antioxidant enzyme B166"  
/protein\_id="AAF03750.1"  
/db\_xref="GI:6103724"  
/translation="MGLAGVCALRRSAGYILVGGAGGSAASAAARRCSEGEWASGV  
SFSSRAAAMAPIKVGDALPAVEFEGEPGNKNVLAELFRGKGVLFVGPAGTGC  
SK  
THLPFGVEQAEALKAGVQVACLSDAFVTCGEMRAHKAEGKVRLLADPTGAFG  
K  
TDLILLDDLSVIFGNRLKRFMSVVDGIVKALNVEPDGTGLTCSLAPNISQL"

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE

99351776  
1042490  
2 (bases 1 to 805)  
Knoops,B., Wattiez,R., Falmagne,P., Hermans,C. and Bernard,A.  
Direct Submission  
Submitted (04-DEC-1998) Department of Biology, Catholic University  
of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium  
Location/Qualifiers  
1. .805  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11q13; about 7cR from marker D11S913; between  
markers D11S1963 and D11S4407"  
/tissue\_type="lung"  
37. .681  
/note="hAOEB166; alkyl hydroperoxide reductase;  
thioredoxin peroxidase; represents prxv peroxiredoxin  
subfamily in mammals"  
/codon\_start=1  
/product="antioxidant enzyme B166"  
/protein\_id="AAF03750.1"  
/db\_xref="GI:6103724"  
/translation="MGLAGVCALRRSAGYILVGGAGGSAASAAARRCSEGEWASGV  
SFSSRAAAMAPIKVGDALPAVEFEGEPGNKNVLAELFRGKGVLFVGPAGTGC  
SK  
THLPFGVEQAEALKAGVQVACLSDAFVTCGEMRAHKAEGKVRLLADPTGAFG  
K  
TDLILLDDLSVIFGNRLKRFMSVVDGIVKALNVEPDGTGLTCSLAPNISQL"

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE

99351776  
1042490  
2 (bases 1 to 805)  
Knoops,B., Wattiez,R., Falmagne,P., Hermans,C. and Bernard,A.  
Direct Submission  
Submitted (04-DEC-1998) Department of Biology, Catholic University  
of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium  
Location/Qualifiers  
1. .805  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11q13; about 7cR from marker D11S913; between  
markers D11S1963 and D11S4407"  
/tissue\_type="lung"  
37. .681  
/note="hAOEB166; alkyl hydroperoxide reductase;  
thioredoxin peroxidase; represents prxv peroxiredoxin  
subfamily in mammals"  
/codon\_start=1  
/product="antioxidant enzyme B166"  
/protein\_id="AAF03750.1"  
/db\_xref="GI:6103724"  
/translation="MGLAGVCALRRSAGYILVGGAGGSAASAAARRCSEGEWASGV  
SFSSRAAAMAPIKVGDALPAVEFEGEPGNKNVLAELFRGKGVLFVGPAGTGC  
SK  
THLPFGVEQAEALKAGVQVACLSDAFVTCGEMRAHKAEGKVRLLADPTGAFG  
K  
TDLILLDDLSVIFGNRLKRFMSVVDGIVKALNVEPDGTGLTCSLAPNISQL"

FEATURES  
source

1. .805  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11q13; about 7cR from marker D11S913; between  
markers D11S1963 and D11S4407"  
/tissue\_type="lung"  
37. .681  
/note="hAOEB166; alkyl hydroperoxide reductase;  
thioredoxin peroxidase; represents prxv peroxiredoxin  
subfamily in mammals"  
/codon\_start=1  
/product="antioxidant enzyme B166"  
/protein\_id="AAF03750.1"  
/db\_xref="GI:6103724"  
/translation="MGLAGVCALRRSAGYILVGGAGGSAASAAARRCSEGEWASGV  
SFSSRAAAMAPIKVGDALPAVEFEGEPGNKNVLAELFRGKGVLFVGPAGTGC  
SK  
THLPFGVEQAEALKAGVQVACLSDAFVTCGEMRAHKAEGKVRLLADPTGAFG  
K  
TDLILLDDLSVIFGNRLKRFMSVVDGIVKALNVEPDGTGLTCSLAPNISQL"

## CDS

1. .805  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11q13; about 7cR from marker D11S913; between  
markers D11S1963 and D11S4407"  
/tissue\_type="lung"  
37. .681  
/note="hAOEB166; alkyl hydroperoxide reductase;  
thioredoxin peroxidase; represents prxv peroxiredoxin  
subfamily in mammals"  
/codon\_start=1  
/product="antioxidant enzyme B166"  
/protein\_id="AAF03750.1"  
/db\_xref="GI:6103724"  
/translation="MGLAGVCALRRSAGYILVGGAGGSAASAAARRCSEGEWASGV  
SFSSRAAAMAPIKVGDALPAVEFEGEPGNKNVLAELFRGKGVLFVGPAGTGC  
SK  
THLPFGVEQAEALKAGVQVACLSDAFVTCGEMRAHKAEGKVRLLADPTGAFG  
K  
TDLILLDDLSVIFGNRLKRFMSVVDGIVKALNVEPDGTGLTCSLAPNISQL"

## misc\_feature

193  
/note="probable alternative translation start site"

BASE COUNT  
ORIGIN

163 a 200 c 276 g 166 t

## alignment\_scores:

Quality: 830.00 Length: 162  
Ratio: 5.123 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-486-167A-2 x AF110731 ..  
Align seg 1/1 to: AF110731 from: 1 to: 805

## 1 MetAlaProLysValGlyAspAlaIleProAlaValGluValPheG1 17

193 ATGCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTGA 242

## 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysLysL 34

243 AGGGACCCAGGGAACAAAGTGAACCTGGCAGAGCTGTTCAAGGCAAGA 292



PUBMED 10514471  
 REFERENCE 2 (bases 1 to 840)  
 AUTHORS Yamashita, H., Avraham, S., London, R., Van Veidhoven, P.P.,  
 Subramani, S. and Avraham, H.  
 TITLE Direct Submission  
 JOURNAL Submitted (31-JAN-1999) Experimental Medicine, Beth Israel  
 Deaconess Medical Center, 4 Blackfan Circle, Boston, MA 02115, USA  
 FEATURES Location/Qualifiers  
 source 1..840  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 CDS 226..714  
 /note="PMP20"  
 /codon\_start=1  
 /product="peroxisomal membrane protein 20"  
 /protein\_id="AAF27531.1"  
 /db\_xref="GI:6746355"  
 /translation="MAPIKVGDAIPAVEVEFEGEPGNKVNLAELFKGKGVLFVGPVAF  
 TPQCKSTHLPGRVEQAEALKAKGVQVACLVSNDVFTGEGRAHKAEGKVRLLADPT  
 GAFGETDLLDDSLVSIFGNRLKRFMSVVDGIVKALNVEPDGTGLTCSLAPNIIS  
 QL"

BASE COUNT 166 a 214 c 289 g 171 t  
 ORIGIN

alignment\_scores:  
 Quality: 830.00 Length: 162  
 Ratio: 5.123 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-486-167A-2 x AF124993

Align seg 1/1 to: AF124993 from: 1 to: 840

1 MetAlaProLleLysValGlyAspAlaIleProAlaValGluValPheG1 17  
 |||||  
 226 ATGCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTGA 275  
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysLysL 34  
 |||||  
 276 AGGGAGCCAGGAAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 325  
 34 ySgLyValLeuPheGlyValProGlyValAlaPheThrProGlyCysSerLys 50  
 |||||  
 326 AGGTGTGCTGTTGGAGTCTCTGGGGCTTCACCCCTGGATGTTCCAAG 375  
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67  
 |||||  
 376 ACACACCTGCCAGGGTTTGTGGAGCAGGCTGAGGCTCTCAAGGCCAAGG 425  
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84  
 |||||  
 426 AGTCCAGGTGGTGGCTGTGAGTGTATGATGCTTTGTGACTGGCG 475  
 84 luTrpGlyArgAlaHisLysAlaGluLysValArgLeuLeuAlaAsp 100  
 |||||  
 476 AGTGGGGCCGAGCCACACAGCGGAGGCAAGTTCCGGCTCTGCTGAT 525  
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117  
 |||||  
 526 CCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCGCT 575  
 117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134  
 |||||  
 576 GTGTCCATCTTTGGGAATCGACGTCCTCAAGAGGTCTCCATGGTGTAC 625  
 134 InAspGlyLleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150  
 |||||  
 626 AGGATGGCATAGTGAAGCCCTGATGTGGACCAAGATGGCACAGCCCTC 675  
 151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162  
 |||||

676 ACCTGACGCTGGCACCACCAATATCATCTCACAGCTC 711

seq\_name: gb\_pr5:AF242525

seq\_documentation\_block:

LOCUS AF242525 875 bp mRNA PRI 15-AUG-2000  
 DEFINITION Homo sapiens hypothetical protein SBB110 mRNA, complete cds.  
 ACCESSION AF242525

VERSION AF242525.1 GI:9802047

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 875)

AUTHORS Zhang, W., Li, N., Wan, T. and Cao, X.

TITLE Direct Submission

JOURNAL Submitted (08-MAR-2000) Department of Immunology, Second Military

Medical University & Shanghai Brilliance Biotechnology Institute,

800 Xiangyin Rd., Shanghai 200433, P.R. China

FEATURES Location/Qualifiers

source

1..875

/organism="Homo sapiens"

/db\_xref="taxon:9606"

36..680

/codon\_start=1

/product="hypothetical protein SBB110"

/protein\_id="AAF99605.1"

/db\_xref="GI:9802048"

/translation="MGLAGVLCALRRSAGYILVGGAGSQAARRCSEGEWASGGVR

SFSRAAAMAPIKVDIPAIVEVEFEGEPGNKVNLAELFKGKGVLFVGPVAFTPGCSK

THLPGFVEQAEALKAKGVQVACLVSNDVFTGEGRAHKAEGKVRLLADPTGAFGKE

TDLLDDSLVSIFGNRLKRFMSVVDGIVKALNVEPDGTGLTCSLAPNIISOL"

BASE COUNT 230 a 206 c 272 g 167 t

ORIGIN

alignment\_scores:

Quality: 830.00 Length: 162  
 Ratio: 5.123 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-486-167A-2 x AF242525

Align seg 1/1 to: AF242525 from: 1 to: 875

1 MetAlaProLleLysValGlyAspAlaIleProAlaValGluValPheG1 17  
 |||||  
 192 ATGCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTGA 241  
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysLysL 34  
 |||||  
 242 AGGGAGCCAGGAAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 291  
 34 ySgLyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
 |||||  
 292 AGGTGTGCTGTTGGAGTCTCTGGGGCTTCACCCCTGGATGTTCCAAG 341  
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67  
 |||||  
 342 ACACACCTGCCAGGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGG 391  
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84  
 |||||  
 392 AGTCCAGGTGGTGGCTGTCTGAGTGTATGATGCTTTGTGACTGGCG 441  
 84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100  
 |||||  
 442 AGTGGGCCGAGCCACACAGCGGAGGCAAGGTTCCGGCTCTCGGCTGAT 491  
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117  
 |||||

492 CCCACTGGGGCCCTTTGGGAAGAGACAGACTTATTACTAGATGATTCGCT 541  
117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134  
542 GGTGTCATCTTTGGGATCGAGCTCAAGAGGTTCTCCATGGTGTAC 591  
134 InAspGlyTleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150  
592 AGGATGGCATAGTGAAGGCCCTGAATGTGAACACGATGGCACAGGCCCTC 641  
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162  
642 ACCCTGAGCTGGCCACCAATATCATCTCACAGCTC 677

seq\_name: gb\_pr4:AF110736  
seq\_documentation\_block: 725 bp mRNA PRI 27-SEP-2000  
LOCUS AF110736 Cercopithecus aethiops peroxiredoxin 5 (PRDX5) mRNA, complete cds.  
DEFINITION AF110736  
ACCESSION AF110736  
VERSION AF110736.2 GI:10305337  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REMARK  
COMMENT  
FEATURES  
source

gene  
CDS  
misc\_feature

1..725  
/organism="Cercopithecus aethiops"  
/db\_xref="taxon:9534"  
/cell\_line="COS-7"  
1..725  
/gene="PRDX5"  
5..652  
/gene="PRDX5"  
/note="AOEB166; PRDX5; alkyl hydroperoxide reductase;  
thioredoxin peroxidase; mitochondrial, peroxisomal, and  
cytosolic"  
/codon\_start=1  
/product="peroxiredoxin 5"  
/protein\_id="AAG13453.2"  
/db\_xref="GI:10305338"  
/translation="MGLAGCVLRRSAGYILGGAARQSVAAATARRRRGGWASGV  
RFSFAAAAMAPIKVGDAIPAVEVEFEGPKNVLAELFKRGKGLVFGVPGFTPCGS  
KTHLPQFVEQALAKAGVQVLACSLDNFVYGEWGRAHKAEGKRVRLADPTGAEGK  
ETDILLDDLSLVIFGNRLKRFSMVYQDGVIVKALNVEPDGTGLTCSLAPSIISQL"  
164  
/gene="PRDX5"

seq\_name: gb\_pr5:AF197952  
seq\_documentation\_block: 780 bp mRNA PRI 29-FEB-2000  
LOCUS AF197952 Homo sapiens thioredoxin peroxidase PMP20 mRNA, complete cds.  
DEFINITION AF197952  
ACCESSION AF197952  
VERSION AF197952.1 GI:6166492  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 780)  
Zhou, Y., Kok, K.H., Chun, A.C., Wong, C.M., Wu, H.W., Lin, M.C.,  
Fung, P.C., Kung, H., and Jin, D.Y.  
Mouse peroxiredoxin V is a thioredoxin peroxidase that inhibits  
p53-induced apoptosis  
Biochem. Biophys. Res. Commun. 268 (3), 921-927 (2000)

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17  
164 ATGGCCCCCATCAAGTGGGAGATGCCATCCCTGCAGTGGAGGTGTTGA 213  
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34  
214 AGGGAGCCAGGCAACAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 263  
34 yGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
264 AGGCTGTGCTGTTGGAGTTCCCGGGGCTTCACGCTGGATGTTCCAG 313  
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67  
314 ACCACCTACAGGGTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAAGG 363  
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84  
364 AGTCCAGCTGTGGCTGTCTGAGTGTAAATGATGCTTTGTGACTGGCG 413  
84 LuTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100  
414 AGTGGGCCCGGACCCACAGGCGGAAGCAAGGTTCGGCTCCTGGCTGAT 463  
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117  
464 CCCACTGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCGTT 513  
117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134  
514 AGTGTCCATCTTTGGGAATCGACGCTCAAGAGGTTCTCCATGGTGTAC 563  
134 InAspGlyTleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150  
564 AGGATGCATAGTGAAGCCCTGAATGTGAACACGATGGCACAGGCCCTC 613  
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162  
614 ACCTGCAGTCTGGCAGCCAGCATCATCTCACAGCTC 649

alignment\_scores:  
Quality: 822.00 Length: 162  
Ratio: 5.074 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 98.765  
alignment\_block:  
US-09-486-167A-2 x AF110736 ..  
Align seg 1/1 to: AF110736 from: 1 to: 725

```

MEDLINE      20145535
REFERENCE    2 (bases 1 to 780)
AUTHORS      Jin,D.-Y.
TITLE        Direct Submission
JOURNAL      Submitted (23-Oct-1999) Institute of Molecular Biology, The
FEATURES     University of Hong Kong, 8 Sassoon Road, Hong Kong
SOURCE       Location/Qualifiers
            1. .780
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="11"
            /map="11q13"
            /map="11q13"
            110..754
            /note="type V peroxidase"
            /codon_start=1
            /product="thioredoxin peroxidase PMP20"
            /protein_id="AAF04856.1"
            /db_xref="GI:6166493"
            /translation="MGLAGVCLRRSAGYILVGGAGGQSAARRCSEGEWASGGVR
            SFSSRAAANAPIKVGDAPAVEFEGEPGNKNLAEFLKGGKGVLFVPGFTPGCSK
            THLPFGVEQAELKAGVQVACLSDAFVTGEMGRATKAEGKVRLLADPTGAFGCK
            TDLLDDSLVSTFGNRRRLKRFMSVVDGIVKALNVEPDGTGLTCSLAPNIIISQL"
BASE COUNT   147 a 203 c 278 g 152 t
ORIGIN
alignment_scores:
  Quality: 820.00      Length: 162
  Ratio: 5.093        Gaps: 0
  Percent Similarity: 99.383 Percent Identity: 99.383
alignment_block:
US-09-486-167A-2 x AF197952 ..
Align seg 1/1 to: AF197952 from: 1 to: 780
1  MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
|||||
266 ATGGCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTGA 315
|||||
17  uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
|||||
316 AGGGAGCCAGGACACAGGTGAACCTGGCAGAGCTGTTCAAGGGCAGA 365
|||||
34  ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
|||||
366 AGGTGTGCTGTTGGAGTCTCTGGGGCTTCAACCCCTGGATGTTCCAAG 415
|||||
51  ThrHisLeuProGlyPheValGluAlaGluAlaLeuLysAlaLysG1 67
|||||
416 ACACACCTGCCAGGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGG 465
|||||
67  yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
|||||
466 AGTCCAGGTGTGGCTGCTGAGCTGTAATGATGCCTTTGTGACTGGCG 515
|||||
84  luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
|||||
516 AGTGGGGCCGACCAACAAGCGGAGCAAGGTTCGGCTCCTGGCTGAT 565
|||||
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117
|||||
566 CCCACTGGGGCTTTGGAGAGGACAGACTTATAGTAGATGATTCGCT 615
|||||
117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
|||||
616 GGTGTCCATCTTTGGGAATCGAGCTCTCAAGAGGTCTCCATGGTGTGAT 665
|||||
134 lnAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
|||||
666 AGGATGCATAGTAGAAGCCCTGAATGTGAACACAGATGGCACAGGCCTC 715
|||||
151 ThrCysSerLeuAlaProAsnIleSerGlnLeu 162

```

```

|||||
716 ACCTGCAGCCTGGCACCAATATCATCTCACAGCTC 751
seq_name: gb_pr4:AF110734
seq_documentation_block:
LOCUS      AF110734          725 bp      mRNA
DEFINITION Papio hamadryas peroxidexin 5 (PRDX5) mRNA, complete cds.
ACCESSION  AF110734
VERSION    AF110734.2  GI:10305333
KEYWORDS
SOURCE     baboon.
ORGANISM   Papio hamadryas
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
            Cercopithecinae; Papio.
REFERENCE  1 (bases 1 to 725)
AUTHORS   Knoops,B., Clippe,A., Bogard,C., Arsalane,K., Wattiez,R.,
            Hermans,C., Duconseille,E., Falmagne,P. and Bernard,A.
TITLE      Cloning and characterization of AOEB166, a novel mammalian
            antioxidant enzyme of the peroxidexin family
JOURNAL    J. Biol. Chem. 274 (43), 30451-30458 (1999)
MEDLINE    99452929
PUBMED     10521424
REFERENCE  2 (bases 1 to 725)
AUTHORS   Knoops,B. and Cherif,H.
TITLE      Cloning and characterization of baboon AOEB166/PRDX5
JOURNAL    Unpublished
REFERENCE  3 (bases 1 to 725)
AUTHORS   Knoops,B., Wattiez,R., Falmagne,P., Hermans,C. and Bernard,A.
TITLE      Direct Submission
JOURNAL    Submitted (04-DEC-1998) Department of Biology, Catholic University
            of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium
REFERENCE  4 (bases 1 to 725)
AUTHORS   Knoops,B., Wattiez,R., Falmagne,P., Hermans,C. and Bernard,A.
TITLE      Direct Submission
JOURNAL    Submitted (26-SEP-2000) Department of Biology, Catholic University
            of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium
REMARK     Sequence update by submitter
COMMENT     On Sep 26, 2000 this sequence version replaced gi:10129958.
FEATURES   Location/Qualifiers
            1..725
            /organism="Papio hamadryas"
            /db_xref="taxon:9557"
            1..725
            /gene="PRDX5"
            5..652
            /gene="PRDX5"
            /note="AOEB166; PRDX5; alkyl hydroperoxide reductase;
            thioredoxin peroxidase; mitochondrial, peroxisomal, and
            cytosolic"
            /codon_start=1
            /product="peroxidexin 5"
            /protein_id="AAG13451.2"
            /db_xref="GI:10305334"
            /translation="MGLAGVCLRRSAGYILGGAAGQSAVAATAARRRRSGGNASGCV
            RFSRAAANAPIKVGDAPAVEFEGEPGNKNLAEFLKGGKGVLFVPGFTPGCS
            KTHLPFGVEQAELKAGVQVACLSDAFVTGEMGRATKAEGKVRLLADPTGAFGCK
            ETDLLDDSLVSTFGNRRRLKRFMSVVDGIVKALNVEPDGTGLTCSLAPNIIISQL"
            164
            /gene="PRDX5"
            /note="probable alternative translation start site"
BASE COUNT 141 a 186 c 246 g 152 t
ORIGIN

```

```

alignment_scores:
  Quality: 818.00      Length: 162
  Ratio: 5.049        Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 98.148
alignment_block:
US-09-486-167A-2 x AF110734 ..

```

Align seg 1/1 to: AF110734 from: 1 to: 725

```

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
|||||
164 ATGGCCCCGATCAAGTGGAGATGCCATCCCTGCAGTGGAGGTGTTGA 213
|||||
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
|||||
214 AGGGAGCCAGGACAAAGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 263
|||||
34 ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
|||||
264 AGGGTGTCTGTTGGAGTTCCTGGGGGCTTCACGCCCTGGATGTTCCAAG 313
|||||
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
|||||
314 ACCCACTACCAGGGTTTGTGGAGCAGCTGAGGCTCTGAAGGCCAAGG 363
|||||
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
|||||
364 AGTCAGGTGTGGCTCTCTGAGTGTAAATGATGCCCTTGTGACTGGCG 413
|||||
84 LuTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
|||||
414 AGTGGGCGGAGCCACAAAGTGGGAAGCAAGGTTCCGCTCGTGCCTGAT 463
|||||
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
|||||
464 CCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCATT 513
|||||
117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
|||||
514 AGTGTCATCTTTGGGAATGACGCTCTCAAGAGTTCTCCATGGTGGTAC 563
|||||
134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
|||||
564 AGGATGGCATAGTAAGGCCCTGAATGTGAACCAAGATGGCACAGGCCCTC 613
|||||
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162
|||||
614 ACCTGCAGTCTGGCACCCCAACATCTCTCACAGCTC 649
|||||

```

seq\_name: gb\_pat1:A99101

```

seq_documentation_block: 675 bp DNA PAT 20-SEP-2000
LOCUS A99101
DEFINITION Sequence 5 from Patent WO9909054.
ACCESSION A99101
VERSION A99101.1 GI:6782051
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 675)
Falmagne, P., Wattiez, R., Bernard, A., Hermans, C. and Knoops, B.
PEROXISOME-ASSOCIATED POLYPEPTIDE, NUCLEOTIDE SEQUENCE ENCODING
said polypeptide and their uses in the diagnosis and/or the
treatment of lung injuries and diseases, and of oxidative
stress-related disorders
JOURNAL Patent: WO 9909054-A 5 25-FEB-1999;
UNIV MONS HAINAUT (BE) ; FALMAGNE PAUL (BE) ; WATTIEZ RUDDY (BE) ;
BERNARD ALFRED (BE) ; HERMANS CEDRIC (BE) ; KNOOPS BERNARD (BE) ;
UNIV LOUVAIN (BE)
FEATURES Location/Qualifiers
source 1..675
/organism="Mus sp."
/db_xref="taxon:10095"
99..588
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB70591.1"

```

```

seq_name: gb_pat1:A99101
seq_documentation_block: 675 bp mRNA ROD 27-SEP-2000
LOCUS AF110733
DEFINITION Mus musculus peroxiredoxin 5 (Prdx5) mRNA, complete cds.
ACCESSION AF110733
VERSION AF110733.1 GI:10129956
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 675)
AUTHORS Knoops, B., Clippe, A., Bogard, C., Arsalane, K., Wattiez, R.,

```

```

/db_xref="GI:6782052"
/translation="MAPIKVGDAIPSEVEFEGBPBKVKVLAELFKGKGKGLFVGVPQAF
TPGCSKTHLPFGVEQAGALKAKQAVVACLVDNFVIEWGRAHOAEKVRLLADPT
GAFKATDLLLLDLSLFLGNRRKRFSWIDNGIYKALNVEPDGTLGTLCSLAPNILS
QL"
BASE COUNT 140 a 174 c 214 g 147 t
ORIGIN

```

```

alignment_scores:
Quality: 767.00 Length: 162
Ratio: 4.854 Gaps: 0
Percent Similarity: 97.531 Percent Identity: 91.358

```

alignment\_block:

US-09-486-167A-2 x A99101 ..

Align seg 1/1 to: A99101 from: 1 to: 675

```

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
|||||
99 ATGGCCCCGATCAAGTGGAGATGCCATCCCTCAGTGGAGGTATTGA 148
|||||
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
|||||
149 AGGGCAACCGGAAAGAGTGAACCTTGGCAGAGCTGTTCAAGGGCAAGA 198
|||||
34 ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
|||||
199 AAGGTGTTTTTGTGGAGTCCCTGGGGCATTTACACCTGGCTGTTCTAAG 248
|||||
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
|||||
249 ACCCACTGGCTGGTGGTGGAGCAAGCTGGAGCTCTGAAGGCTAAGG 298
|||||
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
|||||
299 AGCGCAGGTGGTGGCTGCTGAGCGTTAATGACGCTCTTTGTGATTGAAG 348
|||||
84 LuTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
|||||
349 AGTGGGTGCGAGCCCAAGGAGGAGGAGGAGGTTCCGCTCGCTGGCTGAC 398
|||||
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
|||||
399 CCCACTGGAGCCTTTGGGAAGGCGACAGACTTATTATTGGATGATCTTT 448
|||||
117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
|||||
449 GGTGCTCTCTTTGGGAATGCTGCGGCTGAAAAGGTTCTCCATGGTGATAG 498
|||||
134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
|||||
499 ACAAGCGCATAGTAGAGGCACTGAAGCTGAGCCAGATGGCACAGGCCCTC 548
|||||
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162
|||||
549 ACCTGCAGCTGGCCCCCAACATCTCTCTCCCAACTC 584
|||||

```

Hermans, C., Duconseille, E., Falmagne, P., and Bernard, A.  
 Cloning and characterization of AOB166, a novel mammalian  
 antioxidant enzyme of the peroxiredoxin family  
 J. Biol. Chem. 274 (43), 30451-30458 (1999)  
 99452929  
 MEDLINE  
 PUBMED  
 10521424  
 REFERENCE  
 2 (bases 1 to 675)  
 AUTHORS  
 Knoops, B., Clippe, A., and Bernard, A.  
 Cloning and characterization of mouse AOB166/PRDX5  
 Unpublished  
 TITLE  
 3 (bases 1 to 675)  
 REFERENCE  
 Knoops, B., Wattiez, R., Falmagne, P., Hermans, C., and Bernard, A.  
 Direct Submission  
 TITLE  
 Submitted (04-DEC-1998) Department of Biology, Catholic University  
 of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium  
 JOURNAL  
 FEATURES  
 Location/Qualifiers  
 1..675  
 source  
 /organism="Mus musculus"  
 /strain="C3H/HeJ"  
 /db\_xref="taxon:10090"  
 /tissue\_type="lung"  
 1..675  
 gene  
 /gene="Prdx5"  
 99..587  
 CDS  
 /gene="Prdx5"  
 /note="AOB166; PRDX5; alkyl hydroperoxide reductase;  
 thioredoxin peroxidase; peroxisomal and cytosolic"  
 /codon\_start=1  
 /product="peroxiredoxin 5"  
 /protein\_id="AAG13450.1"  
 /db\_xref="GI:10129957"  
 /translation="MAPIKVGDAIPSEVEFEGEPGKKNVLAELFKGKGVLFPGVGF  
 TPQSKTHLPFGVQAGALKAKGAVVACLSVNDVFIEMGRAHQAEKVRLLADPT  
 GAFKATDLLDDLSVLFGRNRRKRFMSVIDNGIVKALNVEPDGTGLTCLAPNLS  
 QL"  
 BASE COUNT 140 a 174 c 214 g 147 t  
 ORIGIN  
 alignment\_scores  
 Quality: 767.00 Length: 162  
 Ratio: 4.854 Gaps: 0  
 Percent Similarity: 97.531 Percent Identity: 91.358  
 alignment\_block  
 US-09-486-167A-2 x AF110733 ..  
 Align seg 1/1 to: AF110733 from: 1 to: 675  
 1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17  
 99 ATGGCCCCGATCAAGGTGGAGATGCCATTCCTCAGTGGAGTATTGA 148  
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyL 34  
 149 AGGGGAACCGGAAAGAGGTGAAGTGGCAGAGCTGTCAAGGGCAAGA 198  
 34 ySgLyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
 199 AAGGTGTTTGTGGAGTCCCTGGGGCATTTACACCTGGCTGTCTAAG 248  
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67  
 249 ACCCACCTGCCTGGCTTTTGGAGCAAGCTGGAGCTCTGAAGGCTAAGG 298  
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGly 84  
 299 ACCGAGGTGGGGCTGTCTGAGCGTTAATGACGCTCTTGTGATTGAAG 348  
 84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100  
 349 AGTGGGGTCGAGCCACCAGCAGCAGAGCAAGGTTTCGGCTCCTGGCTGAC 398

101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117  
 399 CCACCTGGAGCCTTTGGGAAGCGGACAGACTTATTATTGGATGATCTTT 448  
 117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134  
 449 GGTGTCTCTCTTTGGGAATCGTCGGCTGAAAGAGTTCTCCATGGTGATAG 498  
 134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150  
 499 ACAACGGCATAGTAGGACCACTGAACGCTGGAGCCAGATGGCAGGCCTC 548  
 151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162  
 549 ACCTGCAGCTGGCCCCCAACATCTCTCCCAACTC 584  
 seq\_name: gb\_rol:AF197951  
 seq\_documentation\_block:  
 LOCUS AF197951 688 bp mRNA ROD 29-FEB-2000  
 DEFINITION Mus musculus thioredoxin peroxidase PMP20 mRNA, complete cds.  
 ACCESSION AF197951  
 VERSION AF197951.1 GI:6166490  
 KEYWORDS  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 REFERENCE 1 (bases 1 to 688)  
 AUTHORS Zhou, Y., Kok, K. H., Chun, A. C., Wong, C. M., Wu, H. W., Lin, M. C.,  
 Fung, P. C., Kung, H., and Jin, D. Y.  
 TITLE Mouse peroxiredoxin V is a thioredoxin peroxidase that inhibits  
 p53-induced apoptosis  
 BIOCHEM. BIOPHYS. RES. COMMUN. 268 (3), 921-927 (2000)  
 JOURNAL 20145535  
 MEDLINE 2 (bases 1 to 688)  
 REFERENCE Jin, D.-Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-OCT-1999) Institute of Molecular Biology, The  
 University of Hong Kong, 8 Sassoon Road, Hong Kong  
 FEATURES  
 Location/Qualifiers  
 source 1..688  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 29..661  
 /note="type V peroxiredoxin"  
 /codon\_start=1  
 /product="thioredoxin peroxidase PMP20"  
 /protein\_id="AA04855.1"  
 /db\_xref="GI:6166491"  
 /translation="MLQLGLRVLGCKASSVLRASVTLGAGRKRKAGWECGSGARSSFS  
 SAVTMAPIKVGDAIPSEVEFEGEPGKKNVLAELFKGKGVLFPGVGFATPGCKSTHLP  
 GFVEQAGALKAKGAVVACLSVNDVFIEMGRAHQAEKVRLLADPTGAGKATDLL  
 LDLSVLSLFGNRRKRFMSVIDNGIVKALNVEPDGTGLTCLAPNLSQL"  
 BASE COUNT 138 a 167 c 233 g 150 t  
 ORIGIN

alignment\_scores  
 Quality: 767.00 Length: 162  
 Ratio: 4.854 Gaps: 0  
 Percent Similarity: 97.531 Percent Identity: 91.358  
 alignment\_block  
 US-09-486-167A-2 x AF197951 ..  
 Align seg 1/1 to: AF197951 from: 1 to: 688  
 1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17  
 173 ATGGCCCCGATCAAGGTGGAGATGCCATTCCTCAGTGGAGTATTGA 222  
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyL 34

```
/db_xref="GI:6103726"
/translation="MVLRFCLVLSIAGSVLRASATWTCVAGRKGAGWEGGARS
FSSAAYTMADIKVCDTRIPSEVEPEGPKVKVLAELPKDKKGLVLFVGPATPCGSKT
HLPGFVEQAGALKAKAGQVYVACLVDNFATVTEWGRAHQAEGKVQLADPTGAFGKET
DLDDSLVSLFNGRRUKRFMSWIDKGVKVALNVEPDTGLTCSLAPNLSQL"
193
misc_feature
193
/note="probable alternative translation start site"
242
variation
242
/note="results in glutamic acid to glycine substitution"
380
/replace="g"
variation
380
/note="results in leucine to proline substitution"
428
/replace="c"
variation
428
/note="results in alanine to valine substitution"
/replace="t"
BASE COUNT 198 a 204 c 256 g 179 t
ORIGIN
alignment_scores:
Quality: 767.00 Length: 162
Ratio: 4.824 Gaps: 0
Percent Similarity: 98.148 Percent Identity: 90.741
alignment_block:
US-09-486-167A-2 x AF110732 ..
Align seg 1/1 to: AF110732 from: 1 to: 837
1 MetAlaProLleLysValGlyAspAlaIleProAlaValGluValPheG1 17
193 ATGCCCCGATCAAGGTGGGAGACACCATTCCTCAGTGGAGGTATTGA 242
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
243 AGGGGAACCTGGAAAGAGGTGAAGTGGAGAGGTCTTCAAGGACAAGA 292
34 ySGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
293 AAGGTGTTTGTGGAGTCCCTGGGGCAATTACACCTGGCTGTTCGAAG 342
51 ThrHisLeuProGlyPheValGluAlaGluAlaLeuLysAlaLysG1 67
343 ACCCATCTGCTGGTGTGGAGCAAGCCGAGCTCTGAAGGCCAAGGG 392
67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
393 AGCACAACTGGTGGCCTGTCTGAGTGTAAATGATGCCTTCGTGACTGCAG 442
84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
443 AGTGGGTCGAGCCACAGCAGCAGAGGAGGTTTCAGCTCCTGGCTGAC 492
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
493 CCCACTGGAGCTTTTGGAAAGAGACAGATTACTACTAGATGATCTTT 542
117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
543 GGTCTCTCTCTTGGGAATCGCTGGCTAAAAAAGGTTCCTCCATGGTATAG 592
134 lnAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
593 ACAAGGGCGTAGTAAGCACTGAACGTGGAGCCGGATGGCAGGCCTC 642
151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
643 ACCTGCAGCCTGGCCCCCAACATCTCTCTCAACAATC 678
seq_name: gb_rol:AF124994
seq_documentation_block:
```

LOCUS AF124994 895 bp mRNA ROD 25-JAN-2000  
 DEFINITION Mus musculus peroxisomal membrane protein 20 mRNA, complete cds.  
 ACCESSION AF124994  
 VERSION AF124994.1 GI:6746356  
 KEYWORDS house mouse.  
 SOURCE Mus musculus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 895)  
 AUTHORS Yamashita, H., Avraham, S., Jiang, S., London, R., Van Veldhoven, P.P., Subramani, S., Rogers, R.A. and Avraham, H.  
 TITLE Characterization of human and murine PMP20 peroxisomal proteins that exhibit antioxidant activity in vitro  
 JOURNAL J. Biol. Chem. 274 (42), 29897-29904 (1999)  
 MEDLINE 99445545  
 PUBMED 10514471  
 REFERENCE 2 (bases 1 to 895)  
 AUTHORS Yamashita, H., Avraham, S., London, R., Van Veldhoven, P.P., Subramani, S. and Avraham, H.  
 TITLE Direct Submission  
 JOURNAL Submitted (31-JAN-1999) Experimental Medicine, Beth Israel Deaconess Medical Center, 4 Blackfan Circle, Boston, MA 02115, USA  
 FEATURES  
 source  
 1..895  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 262..750  
 /note="PMP20"  
 /codon\_start=1  
 /product="peroxisomal membrane protein 20"  
 /protein\_id="AAF27532.1"  
 /db\_xref="GI:6746356"  
 /translation="MAPIKVGDALPSVEFEGEPGKKNLAELFKGKGVLFVGPAGF  
 TPQSKTHLPFGVQAGALKAGVAVCLSVNDVFVIEWGRAHQAEGKVRLLADPT  
 GAFGKATDLLLDSLSLFGNRLKRFMSWIDNGIVKALNVEPDGTLTCSLAPNLS  
 QL"  
 BASE COUNT 181 a 232 c 284 g 198 t  
 ORIGIN  
 alignment\_scores  
 Quality: 767.00 Length: 162  
 Ratio: 4.854 Gaps: 0  
 Percent Similarity: 97.531 Percent Identity: 91.358  
 alignment\_block  
 US-09-486-167A-2 x AF124994 ..  
 Align seg 1/1 to: AF124994 from: 1 to: 895  
 1 MetAlaProIleIysValGlyAspAlaIleProAlaValGluValPheG1 17  
 |||||||  
 262 ATGGCCCCCATCAAGTGGGAGATGCCATTCCTCAGTGGAGGTATTGA 311  
 |||||||  
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyL 34  
 |||||||  
 312 AGGGGAACCGGAAGAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 361  
 |||||||  
 34 ySGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50  
 |||||||  
 362 AAGGTGTTTGTGGAGTCCTGGGGCATTTACACCTGGCTGTTCTAAG 411  
 |||||||  
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67  
 |||||||  
 412 ACCCACTGCTCCCTGGGTTTGTGGAGCAAGCTGGAGCTCTGAAGGCCAAGG 461  
 |||||||  
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84  
 |||||||  
 462 ACCGCAGGTGGTGGCTCTCTGAGGCTTAATGACGCTTTGTGATTGAAG 511  
 |||||||  
 84 luTTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100  
 |||||||

512 AGTGGGGTTCGAGCCACCACGAGGCAAGGTTTCGGCTCCTGGCTGAC 561  
 |||||||  
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117  
 |||||||  
 562 CCCACTGGAGCCTTTGGGAAGCGCAGACTTATTATTGGATGATCTTT 611  
 |||||||  
 117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValG 134  
 |||||||  
 612 GGTGTCCTCTTTGGGAATCGTCGGCTGAAAAGGTTCTCCATGGTGATAG 661  
 |||||||  
 134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150  
 |||||||  
 662 ACAACGGCATAGTAGAAGCACTGAACGTGGAGCCAGATGGCACAGGCCTC 711  
 |||||||  
 151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162  
 |||||||  
 712 ACCTGCGAGCTGGCCCCCAACATCTCTCTCAACTC 747

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2001, 10:44:35 ; Search time 58.22 Seconds  
(without alignments)  
2617.579 Million cell updates/sec

Title: US-09-486-167A-1

Perfect score: 805

Sequence: 1 gccaggagcggagtgaag.....ttgtggtttgcggaaaaaaa 805

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*

- 1: /cgn2.6/prodata/2/ina/5A\_COMB.seq:\*
- 2: /cgn2.6/prodata/2/ina/5B\_COMB.seq:\*
- 3: /cgn2.6/prodata/2/ina/6A\_COMB.seq:\*
- 4: /cgn2.6/prodata/2/ina/6B\_COMB.seq:\*
- 5: /cgn2.6/prodata/2/ina/pctus\_COMB.seq:\*
- 6: /cgn2.6/prodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	798	99.1	993	4	US-08-959-004-2
2	49.4	6.1	7218	1	US-08-232-463-14
3	39.4	4.9	4616	1	US-08-340-203A-1
4	39.4	4.9	4616	2	US-08-452-567-1
5	39.4	4.9	4616	2	US-08-452-567-1
6	39.4	4.9	4616	3	US-09-085-407-1
7	37.2	4.6	4112	1	US-08-340-203A-2
8	37.2	4.6	4112	2	US-08-452-567-2
9	37.2	4.6	4112	2	US-08-452-567-2
10	37.2	4.6	4112	3	US-09-085-407-2
11	36.2	4.5	68750	3	US-09-335-409-1
12	36	4.5	458	3	US-09-141-000-4
13	34.8	4.3	418	3	US-09-141-000-6
14	34	4.2	4252	2	US-08-475-844-4
15	34	4.2	4252	5	PCT-US95-08429-4
16	33.6	4.2	2219	3	US-08-510-646B-17
17	33.6	4.2	6453	1	US-08-306-691B-14
18	33.6	4.2	6453	3	US-09-209-668-10
19	33.6	4.2	6453	3	US-09-356-952-8
20	33.4	4.1	1611	2	US-08-650-598-1
21	33.4	4.1	35060	3	US-08-814-095-7
22	33.2	4.1	2329	1	US-08-457-797A-8
23	33.2	4.1	2329	1	US-08-812-025-8
24	33.2	4.1	2721	6	5215881-2
25	33.2	4.1	8438	1	US-07-945-283-1
26	32.8	4.1	1866	4	US-09-318-448-33
27	32.8	4.1	2384	1	US-07-814-964-10

28	32.8	4.1	2384	1	US-08-258-442-10
29	32.8	4.1	2384	1	US-08-328-809-5
30	32.8	4.1	2384	5	PCT-US92-11107-10
31	32.6	4.0	2031	2	US-08-933-750C-53
32	32.6	4.0	2031	3	US-09-234-613-53
33	32.4	4.0	12001	1	US-08-458-568A-11
34	32.2	4.0	1528	4	US-09-362-473-13
35	32.2	4.0	2580	3	US-09-050-863-2
36	32.2	4.0	5452	2	US-09-130-114-1
37	32.2	4.0	9600	4	US-08-910-647-1
38	32.2	4.0	10596	1	US-07-884-811-15
39	32.2	4.0	10596	1	US-07-885-971-15
40	32.2	4.0	10596	1	US-08-087-783A-15
41	32.2	4.0	10596	1	US-08-194-087B-15
42	32.2	4.0	10596	2	US-08-194-087-15
43	32.2	4.0	10596	5	PCT-US93-04648-15
44	32	4.0	320	4	US-09-165-264-11
45	31.6	3.9	319	4	US-09-165-264-8

#### ALIGNMENTS

#### RESULT 1

US-08-959-004-2  
; Sequence 2, Application US/08959004  
; Patent No. 6197543

#### GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Yue, Henry  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Shah, Purvi  
; APPLICANT: Kaser, Matthew

; TITLE OF INVENTION: HUMAN VESICLE MEMBRANE PROTEIN-LIKE

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

#### COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/959,004  
; FILING DATE: Herewith  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:

#### FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0414 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-855-0555

; TELEFAX: 650-845-4166

#### TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 993 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: BRAITUT01

; CLONE: 743725





Mon Oct 22 14:46:53 2001

APPLICANT: Baylin, Stephen B.  
APPLICANT: Wales, Michele M.  
TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE, HIC-1  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: California  
COUNTRY: USA  
ZIP: 92037

	Query Match	4.9%;	Score 39.4;	DB 2;	Length 4616;	
	Best Local Similarity	48.4%;	Pred. No. 0.091;			
	Matches 139;	Conservative	0;	Mismatches 146;	Indels	Gaps
y	1	gccagaggcgagtggaagtggcgtgsggcggggtatgggactagctggcggtgcgcc	60			
b	1575	CGTGGGCGCCGCAGGCCCGCGCCGACCATAGSGGCGTAGCGCGCGCGCGC	1516			
y	61	ctgagacgtctcagcgggctatatctctcgttgggtggccggcggtcagtcctgcggcacgc	120			
b	1515	CGCGCGCGCCCCGCAGGTGGCGAGTACTTCCTGGCGCGTTGAGGCGGTTCTTGACACGC	1456			
y	121	gcagcaagacggtgcagtgaaagagagtg--ggcgtctggcggggtccgcagatttcagcaa	178			
b	1455	CCACGAGGTGCGGGATCTCGAGTAGCTGGCGGCGSCCAGCACGGCGCCACAGCTCGGCT	1396			
y	179	gagccgctgcagccaatcaagttgggagaatgccatcccagcagtggaaggtgt	238			
b	1395	CAGCCCCGGGGCGACGGCCGCGCGCCGACGCGCCCTCTGCGCGTGCAGCAGCGGCCGG	1336			
y	239	ttagagggggcagggacaaggttgaaacctggcagagctgttcaag	285			
b	1395	TCTAATCAATGAGCTCCAGTACTCAGCGGAACACGCCGGGCTCACCATTG	1289			

us-09-486-167a-1.rni

Patent No. 6103877  
GENERAL INFORMATION:  
APPLICANT: Baylin, Stephen B.  
APPLICANT: Wales, Michele M.  
TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE, HIC-1  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4325 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: California  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/085,407  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/340,203  
FILING DATE: 15-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Ph.D., Lisa A.  
REGISTRATION NUMBER: P-38,347  
REFERENCE/DOCKET NUMBER: 07265/039001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 455-5100  
TELEFAX: (619) 455-5110  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4616 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
IMMEDIATE SOURCE:  
CLONE: HIC-1 polynucleotide  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..4616  
US-09-085-407-1

Query Match	4.9%	Score 39.4	DB 3	Length 4616
Best Local Similarity	48.4%	Pred. No. 0.091		
Matches 139	Conservative	0	Mismatches 146	Indels
QY 1	gccaggaggcgagtggaagtggcgtggggcggggtatgggaactagctggcgctgtgcgcc	60		
Db				
1575	CGTGTGGCGGCCGAGGCCCGCGCGCCGACCATAGGCGCGGTAGCGCGCGCGCGCGC	1516		
QY 61	ctgagaagctcagcggggtctatactctcagttggggcgggcggtcagcttcgcgcgacg	120		
Db				
1515	CGCGCGCCCGCCGAGGTGGCAGTACTTGCCTGTGGCGCTTGAGGCGTTCCTTGACACGG	1456		
QY 121	cgagcaagacggtgcagtgaaaggadagtg -ggcgctctggcggggtccgcagtttcagca	178		
Db				
1455	CCACNGCTCGGGATCTGCAGGTAGCTGGCGCGCGCCAGCACGCGCGCCAGCGTCGGCT	1396		
QY 179	gagcgcgtgagacatgcccccaatcaaggtggagatgccatccagcagtggaaggtgt	238		
Db				
1395	CAGCCCCCGGGCCACGGCGCGCGCGCGAGCCGCTCTGTGCGCGTCAGCCAGCGCGCGG	1336		
QY 239	tigaaggggagccaggaaacaggtgaacctggcagcgtgttcaag	285		
1335	TCTAGATGAAGTCCAGCACACGACGGCGGAACACGCGCGGGTCCACCATG	1289		
Db				

```

Query Match      4.6%; Score 37.2; DB 1; Length 4112;
Best Local Similarity 50.4%; Pred. No. 0.36;
Matches 117; Conservative 0; Mismatches 113; Indels 2; Gaps 1;

QY      56  ggcgcctgagacgctcagcggctatatactcgtctggcggcggcgtcgagctcgcg 115
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1520 GCGCGCGCGCGCGCGCCCGACAGGTGGCAGTACTTCCCTGGCGCTTGAGGCGTTTCTTGA 1461

QY      116  cagcgcgacgaacgctgctgagtgaaagagatg-ggcgtctggcggggtccgcagttt 173
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1460 CAGCGCCACGAGGTGCGGGATCTCGAGGTAGCTTGGGGGGCGCCAGACGCGCCAGGCT 1401

QY      174  cagcagagccgctgcagccatgcccacaatcaaggTggagatgccatcccagcagtgga 233
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1400 CGGCTCAGCGCCCGGGGCGAGCGCGCGCGCGCGCTCTGCGCGCTCAGCCAGGCG 1341

QY      234  ggtgttttaaggaggacagggaacaaaggTgaacctggcagagctgtttcaag 285
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1340 GCCGTGTAGTAGTAAGTCCAGCACACGAGCGCGGAACAGCGCGGGGCTCACCATG 1289

RESULT      8
US-08-452-567-2/c
; Sequence 2, Application US/08452567
; Patent No. 5846712
; GENERAL INFORMATION:
; APPLICANT: Baylin, Stephen B.

```

```

Query Match          4.6%; Score 37.2; DB 2; Length 4112;
Best Local Similarity 50.4%; Pred. No. 0.36;
Matches 117; Conservative 0; Mismatches 113; Indels 2; Gaps 1;

QY 56 ggcgcctgagacgtcagcggtctatatctcgtggggccggcggtcagctcgcgg 115
Db 1520 GCCGCGCCGCCGCCCGCAGGTGGCAGTACTTGCCTGGCTTGAGGGCTTCTTGGCA 1461
QY 116 cagcggcagcaagaagcggtcagtggaaggagatg--ggcctctggtgggggtccgcagttt 173
Db 1460 CAGCGCCACAGAGTTCGGGGATCTGCAGTACTTGGCGGCCGCCACGCGGCCAGGCT 1401
QY 174 cagcagagccgctgcagcgaatgccccaatacaaggTgggagatgccatccacacagtga 233
Db 1400 CGGCTCAGCCCGCGGGGCCACGCCGCCGCGCTCTTGCCTCTTGCACGAGCG 1341
QY 234 ggtgtttgaagggaagccagggaacaaggtgaacctggcagagctgttcaag 285
Db 1340 GCCGGTGTAGATGAAGTCCAGCACACAGGGCGGAACACGGCGGGCTCACCATG 1289

RESULT 9
US-08-452-427-2/c
; Sequence 2, Application US/08452427
; Patent No. 5922590
; GENERAL INFORMATION:
; APPLICANT: Baylin, Stephen B.
; APPLICANT: Wales, Michele M.
; TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE, HIC-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.

```

STATE: California  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: /09/085,407  
APPLICATION NUMBER: US/09/085,407  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/340,203  
FILING DATE: 15-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Halle, Ph.D., Lisa A.  
REGISTRATION NUMBER: P-38,347  
REFERENCE/DOCKET NUMBER: 07265/039001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 455-5100  
TELEFAX: (619) 455-5110  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4112 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
IMMEDIATE SOURCE:  
CLONE: HIC-1 coding polynucleotide  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1086..2726  
85-09-085-407-2

[illegible]

```

RESULT 11
US-09-335-409-1
; Sequence 1, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335.409
; CURRENT FILING DATE: 1999-06-17

```

Query Match	4.5%	Score 36;	DB 3;	Length 458;
Best Local Similarity	8.1%	Pred. No. 0.32;		
Matches 36;	Conservative 152;	Mismatches 253;	Indels 5;	Gaps 1;
QY	56	gcgcctcgtgagcgtcagcgcggtatatactcgtcgtggcgccgcgcgtcagctgcgg	115	
Db	2	DSVELCUPESFLHYEELLCRMSNDRHIDSSCSFIFKTEPSSPASITDSVNHHSPEG	61	
QY	116	cagcgcgcagcaagcgtgcagtgaaaggagatggcgtctcgtgcgggtccgcagttcca	175	
Db	62	SDASSSYSTNMHGHLQDLPPIYPGAPILGSGFVRKLYDDCSSTIVEDPQTKCEYMLN	121	
QY	176	gcagcgcgcgtgcagccatgcgcgcccaatcaagtggtggagatgccatcccacagcgg	232	
Db	122	SMPRKCLVCGDIIASGHYGVASCEACAKFRKRTTQGNIEYSCPATNECEITRRKRSCQ	181	
QY	233	--aggtgtttgaaggcgccagggaaacaagtggaacctggcagcgtgttccaagggcaa	290	
Db	182	ACRFMKCLVKGMUKGVRIDRVRGGRQYKRRIDAENSPLYNPOLVQPAKKPNKIVSHL	241	
QY	291	gaagggtgtcgtgtttgttgagttcctcgtggcgcttcaccccttgatgttccaagacac	350	
Db	242	LVAEPEKIYAMPDPTVPDSDIKALTTICLDADRELWIIIGWAKHIGPFTSLIADOMSL	301	

```

Query Match          4.3%; Score 34.8; DB 3; Length 418;
Best Local Similarity 8.6%; Pred.No.0.67;
Matches 35; Conservative 139; Mismatches 228; Indels 5; Gaps 1;

QY 56 ggcgcctgagacgtcagcgggctatatactctcggtggggccgcggcagtcgtcg 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 DSVELCLPESFSLHYEEELLCRMKNKDRHIDSSCSFLTKTFSPSPASLTSVNHHS 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 116 cagcgcgcagcaagcagtcgagtaagaagagtgggcgctctggcggggtccgcagttca 175
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 SDAGSSYSSTMWNGHQLDGPLYPSPAPILGGSGPVRKLYDDCSSTIVEDPQTKEYMLN 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 176 gcagagcgcgtgcagccatgcccacatcaagtcgagtgagatgccatcccaagcagtg 232
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 SMPKRLVLGCGDIASGVHYGVASCEACKAFKFTIOGNIIEYCPATNECEBITKRRKSCQ 181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 233 --aggtgtttgaaggggagccaggaacaggtgaaacctggcagagctgttcaaggcgca 290
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 ACFMKCLKVGMLEKGVRLDRVGGKQYKRRIDAENSPVLPQLVQPAKKPYNKIVSHL 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 291 gaagggtgtgctgtttggagtctctgggccttcacccctggatgtccaagacacact 350
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 242 LVAEPEKIAMPDPPTVPSDIAKLTLCDLADRELVIIGWAKHIFGFSTLSLADQMSLL 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 351 gccagggtttggagcagcgtgagcgtctgagcgcgaaggagcaggtccaggtggtgcctg 410
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 302 QSAMWEILITLVVYVYSLSFDEFLVADDDYIMBDDQSKLAGLLDLNNAIILQVKKYSMKL 361
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 411 tctgagtgtaatgatgcctttgtactgagcagtgggcgccagccc 457
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 362 EKEEFVTLKAIALANSDSMHIEDVEAVQKLDVLHQLQDYEAGQH 408
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
US-08-475-844-4/c
; Sequence 4, Application US/08475844
; Patent No. 5972643
; GENERAL INFORMATION:
; APPLICANT: Lobanenkov, Victor V.

```

us-09-486-167a-1.rni

Mon Oct 22 14:46:53 2001

APPLICANT: Neiman, Paul E.  
APPLICANT: Klenova, Elena M.  
APPLICANT: Goodwin, Graham H.  
APPLICANT: Filippova, Galina N.  
APPLICANT: Collins, Steven J.  
TITLE OF INVENTION: CTCF  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: One Market Plaza  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,844  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/261,680  
FILING DATE: 17-JUN-1994  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 14538A-11-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-467-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4252 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
ORIGINAL SOURCE:  
ORGANISM: Gallus domesticus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 679..2865  
US-08-475-844-4

	Query Match	4.2%;	Score 34;	DB 2;	Length 4252;	
	Best Local Similarity	45.5%;	Pred. NO. 2.9;	Mismatches 145;	Indels	Gaps
	Matches 121;	Conservative	0;			0;
QY	5	gaaggcggagtggaagtggccgtggaacgggtatgggaactagctggcgctgtgcgccctga	64			
Db	491	GGCGGCGGGCGGGAGTGCCTTCGCGAGCGCGGTGCCCGACGCCCGCCGCAAGA	432			
QY	65	gacgctcagcgggctatatctcgttgttggggccgcgcgtcaacttcgcgcagcgacg	124			
Db	431	CGGGGAGGCGAGGGGGGAGGGACCGCGCGCGGGGAAGCGCGCGCGGGGCAT	372			
QY	125	caagacgggtcagtggaaggagagtggcgctgttgcgggggtccgcagtttcagagaccg	184			
Db	371	CTRGGGAGGCGGGCGGGGAGGGGGCGCGGGCGCGGCCAACATCACAACCGGGG	312			
QY	185	ctgcagccatggccccaatcaagtggagatgccatcccagcagtgagggtgtttgaag	244			
Db	311	CTGGGGCGGGCGGGTGCTACTTCGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG	252			
QY	245	gggagccagggaacaagtgtaacctg	270			
Db	251	GGCGGCGCGCGCTGAGGGGAGGCTTG	226			

search completed: October 22, 2001, 12:13:46  
Job time: 5351 sec

```

15
RESULT
PCT-US95-08429-4/C
; Sequence 4, Application PC/TUS9508429
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: CTCF
; NUMBER OF SEQUENCES: 21
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08429
; FILING DATE: 15-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/261,680
; FILING DATE: 17-JUN-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14538A-11-1PC
; TELEPHONE: 206-467-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; ORIGINAL SOURCE:
; ORGANISM: Gallus domesticus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 679..2865
;
PCT-US95-08429-4

```

Query Match	4.2%;	Score 34;	DB 5;	Length 4252;	
Best Local Similarity	45.5%;	Pred. No. 2.9;	Mismatches 145;	Indels 0;	Gaps 0;
Matches 121;	Conservative				
Y	5	gaagcgagtgaaagtgcctggtgggcgggtatgtagactagctggtgtgcccctga	64		
b	491	GGCGGCGGGCGGGCGGAGTGCCTGCGAGCGCGGTGCGCGGACCGCCGCGACGA	432		
Y	65	gacgtctacggggctatactctgctgtgggcggcggtcagctctgcgcagcggcag	124		
b	431	GCGGGAGGGGAGGGGGGAGGGGACGGCGCGCGGGGAAGCGCGCGCGGGGCAT	372		
Y	125	caagacggttcagtagaagagagtgggcgtctctggcgggtctgcgcagtttcagcagcgcg	184		
b	371	GCTGGGAGGCGCGGGCGGGGAGGGGGGCGCGGGCGCGCCACTTCAAACGCGGG	312		
Y	185	ctgcagccattgcccacaatcaagttggagatgccatccacagcagtggaggtgtttgaag	244		
b	311	CTGGGGCGGCGGTGCATTCGCGTGCCTGCGTGCCTGCACGTGGTAGGAGGGCGG	252		
Y	245	gggagccagggaacaaggtgaacctg	270		
b	251	GGCGGCGCGCGCTGAGGGGAGGCTG	226		

search completed: October 22, 2001, 12:13:46  
Job time: 5351 sec

Mon Oct 22 14:46:53 2001

us-09-486-167a-1.rni

Page 9

